1915–1916

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

ORMOND STONE | PIERRE BOUTROUX
L. P. EISENHART | T. H. GRONWALL
OSWALD VEBLEN | J. H. M. WEDDERBURN

WITH THE COÖPERATION OF

J. W. ALEXANDER | A. A. BENNETT
H. H. MITCHELL

PUBLISHED UNDER THE AUSPICES OF PRINCETON UNIVERSITY

SECOND SERIES, VOL. 17

LANCASTER, PA., AND PRINCETON, N. J.

N.

PRESS OF
THE NEW ERA PRINTING COMPANY
LANCASTER, PA.

INDEX

PAGE

ALEXANDER, J. W., Functions which Map the Interior of the Unit Circle upon Simple Regions . . . 12
BEAL, F. W., A Congruence of Circles . . . 180
BENNETT, ALBERT A., The Iteration of Functions of One Variable . . . 23
BENNETT, ALBERT A., Note on an Operation of the Third Grade . . . 74
BENNETT, ALBERT A., Note on the Preceding Paper . . . 123
BENNETT, ALBERT A., A Case of Iteration in Several Variables . . . 188
BÔCHER, M., On the Wronskian Test for Linear Dependence . . . 167
COBLE, A. B., An Isomorphism between Theta Characteristics and the $(2p + 2)$-point . . . 101
DINES, L. L., A Characteristic Property of Self-Projective Curves . . . 213
EISENHART, L. P., Surfaces with Isothermal Representation of Their Lines of Curvature as Envelopes of Rolling . . . 64
GILLESPIE, D. C., The Cauchy Definition of a Definite Integral . . . 61
GRONWALL, T. H., A Functional Equation in the Kinetic Theory of Gases . . . 1
GRONWALL, T. H., Determination of All Triply Orthogonal Systems Containing a Family of Minimal Surfaces . . . 76
GRONWALL, T. H., (see Jensen) . . . 124
JACKSON, D., Non-essential Singularities of Functions of Several Variables . . . 172
JENSEN, J. L. W. V., An Elementary Exposition of the Theory of the Gamma Function. (Authorized translation from the Danish by T. H. Gronwall.) 124
LEFSCHETZ, S. F., The Arithmetic Genus of an Algebraic Manifold Immersed in Another . . . 197
LONGLEY, W. R., Note on a Theorem on Envelopes . . . 169
POUSSIN, C. DE LA VALLÉE, Demonstration simplifiée du théorème fondamental de M. Montel sur les familles normales de fonctions . . . 5
RITT, J. F., On Certain Real Solutions of Babbage's Functional Equation . . . 113
WILLIAMS, K. P., A Theorem Concerning Real Functions . . . 72

A FUNCTIONAL EQUATION IN THE KINETIC THEORY OF GASES.

By T. H. Gronwall.

Let ξ, η, ζ and ξ_1, η_1, ζ_1 be the velocity components of two spherical molecules at the instant of their collision, α, β, γ the direction cosines of their center line at this instant, and finally ξ', η', ζ' and ξ_1', η_1', ζ_1' the velocity components after the collision. Writing

$$W = \alpha(\xi_1 - \xi) + \beta(\eta_1 - \eta) + \gamma(\zeta_1 - \zeta),$$

we then have

$$\xi' = \xi + \alpha W, \quad \eta' = \eta + \beta W, \quad \zeta' = \zeta + \gamma W,$$
$$\xi_1' = \xi_1 - \alpha W, \quad \eta_1' = \eta_1 - \beta W, \quad \zeta_1' = \zeta_1 - \gamma W,$$

and a complete system of invariants for this linear transformation is given by the four expressions

$$\xi + \xi_1, \quad \eta + \eta_1, \quad \zeta + \zeta_1, \quad \xi^2 + \eta^2 + \zeta^2 + \xi_1^2 + \eta_1^2 + \zeta_1^2.$$

Denoting by $\varphi(\xi, \eta, \zeta)$ the logarithm of the function defining the distribution of velocities, it may be shown that $\varphi(\xi, \eta, \zeta) + \varphi(\xi_1, \eta_1, \zeta_1)$ is invariant under the linear transformation in question,* and consequently

$$(1) \quad \varphi(\xi, \eta, \zeta) + \varphi(\xi_1, \eta_1, \zeta_1) = f(\xi + \xi_1, \eta + \eta_1, \zeta + \zeta_1, \xi^2 + \eta^2 + \zeta^2 + \xi_1^2 + \eta_1^2 + \zeta_1^2).$$

Assuming the existence of all six partial derivatives of the second order of φ, it is readily seen that the general solution of this functional equation is†

$$(2) \qquad \varphi(\xi, \eta, \zeta) = a + b_1\xi + b_2\eta + b_3\zeta + c(\xi^2 + \eta^2 + \zeta^2)$$

with constant coefficients.

It is the purpose of the present note to prove this result without any other assumption than that of the continuity of φ for all finite values of the variables.

From (1) subtract the equation obtained by setting all the variables

* See, for a proof involving a minimum of assumptions, Hilbert, "Begründung der kinetischen Gastheorie," Math. Annalen, vol. 72, reprinted in his Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, pp. 267–282.

† Boltzmann, Vorlesungen über Gastheorie, vol. 1, pp. 128–131.

equal to zero; with the notations

$$(3)\qquad \begin{aligned} \varphi_1(\xi, \eta, \zeta) &= \varphi(\xi, \eta, \zeta) - \varphi(0, 0, 0) = \varphi(\xi, \eta, \zeta) - a, \\ f_1(\xi + \xi_1, \cdots) &= f(\xi + \xi_1, \cdots) - f(0, 0, 0, 0) \end{aligned}$$

we obtain

$$(4)\qquad \begin{aligned} &\varphi_1(\xi, \eta, \zeta) + \varphi_1(\xi_1, \eta_1, \zeta_1) \\ &\qquad = f_1(\xi + \xi_1, \eta + \eta_1, \zeta + \zeta_1, \xi^2 + \eta^2 + \zeta^2 + \xi_1^2 + \eta_1^2 + \zeta_1^2), \end{aligned}$$

$$(5)\qquad \varphi_1(0, 0, 0) = 0, \quad f_1(0, 0, 0, 0) = 0.$$

Writing (4) for the two sets of arguments $\xi + \xi_1$, $\eta + \eta_1$, $\zeta + \zeta_1$ and 0, 0, 0, we find with the aid of (5)

$$\begin{aligned} &\varphi_1(\xi + \xi_1, \eta + \eta_1, \zeta + \zeta_1) \\ &\qquad = f_1(\xi + \xi_1, \eta + \eta_1, \zeta + \zeta_1, (\xi + \xi_1)^2 + (\eta + \eta_1)^2 + (\zeta + \zeta_1)^2). \end{aligned}$$

In order to make the fourth argument in f_1 coincide with that in (4), we assume that

$$\xi\xi_1 + \eta\eta_1 + \zeta\zeta_1 = 0,$$

and under this condition we have

$$\varphi_1(\xi + \xi_1, \eta + \eta_1, \zeta + \zeta_1) = \varphi_1(\xi, \eta, \zeta) + \varphi_1(\xi_1, \eta_1, \zeta_1).$$

Our condition is fulfilled by making $\xi_1 = \eta = \zeta = 0$, whence

$$\varphi_1(\xi, \eta_1, \zeta_1) = \varphi_1(\xi, 0, 0) + \varphi_1(0, \eta_1, \zeta_1),$$

and also by making $\xi = \xi_1 = \eta_1 = \zeta = 0$, which gives

$$\varphi_1(0, \eta, \zeta_1) = \varphi_1(0, \eta, 0) + \varphi_1(0, 0, \zeta_1).$$

Replacing η_1 and ζ_1 by η and ζ in the last two equations, we therefore obtain

$$(6)\qquad \varphi_1(\xi, \eta, \zeta) = \varphi_1(\xi, 0, 0) + \varphi_1(0, \eta, 0) + \varphi_1(0, 0, \zeta).$$

Combining (4) and (6), we see at once that

$$(7)\qquad \begin{aligned} &f_1(\xi + \xi_1, \eta + \eta_1, \zeta + \zeta_1, \xi^2 + \eta^2 + \zeta^2 + \xi_1^2 + \eta_1^2 + \zeta_1^2) \\ &\qquad = f_1(\xi + \xi_1, 0, 0, \xi^2 + \xi_1^2) + f_1(0, \eta + \eta_1, 0, \eta^2 + \eta_1^2) \\ &\qquad\qquad + f_1(0, 0, \zeta + \zeta_1, \zeta^2 + \zeta_1^2). \end{aligned}$$

We introduce the notations

$$(8)\qquad \begin{aligned} \xi + \xi_1 &= x, \quad \eta + \eta_1 = y, \quad \zeta + \zeta_1 = z, \\ \xi^2 + \xi_1^2 &= u, \quad \eta^2 + \eta_1^2 = v, \quad \zeta^2 + \zeta_1^2 = w; \end{aligned}$$

these six quantities are independent variables which, however, to insure

the reality of ξ, ξ_1, etc., must satisfy the inequalities

$$x^2 \leqq 2u, \quad y^2 \leqq 2v, \quad z^2 \leqq 2w. \tag{9}$$

Now (7) may be written in the form

$$f_1(x, y, z, u + v + w) = f(x, u) + g(y, v) + h(z, w), \tag{10}$$

where f, g and h are continuous in the region (9) and vanish for zero values of the variables.

Permutating u and v, which is permissible if x^2 and y^2 are both less than or equal to the smaller of $2u$ and $2v$, we find

$$f(x, u) + g(y, v) = f(x, v) + g(y, u), \tag{11}$$

and on making $y = 0$, $v = v_0 =$ const., the last equation gives

$$f(x, u) = f(x) + \rho(u) \tag{12}$$

for $x^2 \leqq 2u$ and $x^2 \leqq 2v_0$. Here, $\rho(u) = g(0, u)$ is continuous for all $u \geqq 0$ and vanishes for $u = 0$, and since (12) may be written $f(x) = f(x, u) - \rho(u)$, it is clear that $f(x)$ is continuous for all finite values of x, and vanishes for $x = 0$. Similarly we obtain $g(y, v) = g(y) + \sigma(v)$, where $g(0) = \sigma(0) = 0$, and substituting in (11) we get $\sigma(v) - \rho(v) = \sigma(u) - \rho(u)$, or making $u = 0$, $\sigma(v) = \rho(v)$. We thus finally have

$$\begin{aligned} f(x, u) &= f(x) + \rho(u), \\ g(y, v) &= g(y) + \rho(v), \\ h(z, w) &= h(z) + \rho(w), \end{aligned} \tag{13}$$

where f, g and h are continuous for all and ρ for positive values of the variable, and all four functions vanish for zero values of their arguments. Substituting in (10) and making $x = y = z = w = 0$, we obtain an equation of the form $\rho_1(u + v) = \rho(u) + \rho(v)$, and making $v = 0$ we see that $\rho_1(u) = \rho(u)$, so that finally

$$\rho(u + v) = \rho(u) + \rho(v). \tag{14}$$

This is the classical functional equation of Cauchy, who showed in his Analyse algébrique that every continuous solution is of the form

$$\rho(u) = cu,$$

where c is a constant.

Consequently, by (8) and (10)

$$\begin{aligned} &f_1(\xi + \xi_1, \eta + \eta_1, \zeta + \zeta_1, \xi^2 + \eta^2 + \zeta^2 + \xi_1^2 + \eta_1^2 + \zeta_1^2) \\ &= f(\xi + \xi_1) + g(\eta + \eta_1) + h(\zeta + \zeta_1) + c(\xi^2 + \eta^2 + \zeta^2 + \xi_1^2 + \eta_1^2 + \zeta_1^2), \end{aligned}$$

and on writing

$$\varphi_2(\xi, \eta, \zeta) = \varphi_1(\xi, \eta, \zeta) - c(\xi^2 + \eta^2 + \zeta^2), \tag{15}$$

equations (4) and (6) become

$$\varphi_2(\xi, \eta, \zeta) + \varphi_2(\xi_1, \eta_1, \zeta_1) = f(\xi + \xi_1) + g(\eta + \eta_1) + h(\zeta + \zeta_1), \tag{4'}$$

$$\varphi_2(\xi, \eta, \zeta) = \varphi_2(\xi, 0, 0) + \varphi_2(0, \eta, 0) + \varphi_2(0, 0, \zeta). \tag{6'}$$

In (4′), make $\eta = \zeta = \xi_1 = \eta_1 = \zeta_1 = 0$, whence $\varphi_2(\xi, 0, 0) = f(\xi)$, and then $\eta = \zeta = \eta_1 = \zeta_1 = 0$, which gives

$$f(\xi + \xi_1) = f(\xi) + f(\xi_1),$$

every continuous solution of which is of the form $f(\xi) = b_1\xi$, where b_1 is a constant. Similarly $\varphi_2(0, \eta, 0) = g(\eta) = b_2\eta$ and $\varphi_2(0, 0, \zeta) = h(\zeta) = b_3\zeta$, so that, by (6′),

$$\varphi_2(\xi, \eta, \zeta) = b_1\xi + b_2\eta + b_3\zeta, \tag{16}$$

and the combination of (16), (15) and (3) finally gives the expression (2) for the most general continuous solution of our functional equation.

PRINCETON UNIVERSITY.

DÉMONSTRATION SIMPLIFIÉE DU THÉORÈME FONDAMENTAL DE M. MONTEL SUR LES FAMILLES NORMALES DE FONCTIONS.

PAR C. DE LA VALLÉE POUSSIN.

1. Dans un Mémoire très intéressant "*Sur les familles de fonctions analytiques qui admettent des valeurs exceptionelles dans un domaine,*" publié dans les Annales de l'Ecole Normale Supérieure, 3ᵉ série t. 29 (1912), M. Montel a démontré un théorème très remarquable, qui fournit peut-être les démonstrations les plus simples des théorèmes de M. Picard sur les propriétès des fonctions uniformes dans le voisinage d'un point critique. La démonstration que M. Montel donne de son théorème fait appel aux propriétés des fonctions modulaires. Il nous a paru intéressant de montrer que les seules propriétés nécessaires peuvent être établies par des considérations complètement indépendantes de la théorie des fonctions elliptiques et de la représentation conforme. Nous nous proposons donc de prouver le théorème en question en nous appuyant uniquement sur les principes les plus généraux de la théorie des fonctions et du prolongement analytique. Cette démonstration trouverait donc facilement sa place dans un Cours général sur la théorie des fonctions analytiques.

2. Nous considérons la fonction $\varphi(x)$ de x complexe définie par l'intégrale, effectuée le long de l'axe réel des u,

$$\varphi(x) = \int_1^\infty \frac{du}{\sqrt{u(u-1)(u-x)}}.$$

La fonction sous le signe d'intégration est *holomorphe* relativement à x tant que x ne se trouve pas sur la ligne d'intégration, qui constitue donc une *coupure*, et $\varphi(x)$ est *holomorphe* hors de la coupure.

Nous fixons la *détermination du radical* en choisissant l'argument θ de $u - x$ entre $\pm\,\pi$. Alors l'argument de $1/\sqrt{u-x}$ est compris entre $\pm\,\pi/2$ et il est de même signe que celui de x (choisi entre $\pm\,\pi$). Avec cette convention, *$\varphi(x)$ a donc sa partie réelle positive et sa partie imaginaire de même signe que celle de x.*

Si x tend vers l'infini d'une manière quelconque, tous les éléments de l'intégrale tendent vers 0 et $\varphi(x)$ également. Mais $\varphi(u)$ n'est nul que si $x = \infty$.

3. Voyons maintenant ce qui arrive sur la coupure. Soit x réel et > 1: c'est un point de la coupure. Désignons par ϵ un infiniment petit positif.

Les deux points infiniment voisins de x de part et d'autre de la coupure sont $x \pm \epsilon i$. Les parties imaginaires et réelles de $\varphi(x \pm \epsilon i)$ s'obtiennent par la décomposition

$$\varphi(x \pm \epsilon i) = \int_1^x \frac{du}{\sqrt{u(u-1)(u-x)}} + \int_x^\infty \frac{du}{\sqrt{u(u-1)(u-x)}}.$$

La première intégrale est purement imaginaire et la seconde réelle. Comme on vient de l'expliquer, la première doit recevoir le signe donné à ϵ et la seconde doit être positive. Ainsi les valeurs de $\varphi(x)$ sont conjuguées de part et d'autre de la coupure.

Les deux intégrales précédentes se ramènent l'une et l'autre à celle qui définit la fonction φ par un changement de variable approprié. En posant

$$u = 1 + \frac{x-1}{v}, \qquad \text{ou} \qquad u = vx,$$

il vient respectivement (x étant réel et > 1)

$$\int_1^x \frac{du}{\sqrt{u(u-1)(u-x)}} = \int_1^\infty \frac{dv}{\sqrt{-v(v-1)(v-1+x)}} = \pm \frac{\varphi(1-x)}{i},$$

$$\int_x^\infty \frac{du}{\sqrt{u(u-1)(u-x)}} = \int_1^\infty \frac{dv}{\sqrt{v(vx-1)(v-1)}} = \frac{\varphi\left(\frac{1}{x}\right)}{\sqrt{x}}.$$

Donc il vient (les signes $\pm$ se correspondant)

$$\varphi(x \pm \epsilon i) = \frac{\varphi\left(\frac{1}{x}\right)}{\sqrt{x}} \pm i\varphi(1-x).$$

Cette formule est très intéressante, car le second membre est holomorphe sur la coupure $1 \text{——} \infty$ sauf aux extrémités. Le second membre fournit donc le prolongement analytique de la fonction $\varphi(x)$ des deux côtés de la coupure. Ainsi la coupure était artificielle et elle ne renferme que deux points critiques 1 et ∞. Toutefois la formule précédente fournit une relation qui se prête mieux à l'étude du prolongement analytique. Elle montre que l'on a

$$\varphi(x + \epsilon i) = \varphi(x - \epsilon i) + 2i\varphi(1-x).$$

Il s'ensuit que, si l'on traverse la coupure, la fonction $\varphi(x)$ se transforme (suivant le sens de la traversée) dans

$$\varphi(x) \pm 2i\varphi(1-x).$$

Cette nouvelle fonction est holomorphe dans le domaine d'uniformité

commun à $\varphi(x)$ et $\varphi(1-x)$, c'est-à-dire dans le domaine limité par les deux coupures $1 —— \infty$ et $0 —— -\infty$, se rapportant respectivement à $\varphi(x)$ et $\varphi(1-x)$. Elle n'a sur la frontière de ce domaine que trois points critiques 0, 1 et ∞.

4. Maintenant le prolongement analytique de la fonction $\varphi(x)$ se poursuit indéfiniment sans nouvelle introduction de coupure, ni rencontre de nouveaux points critiques. Après un nombre quelconque de traversées des deux coupures précédentes, l'expression de $\varphi(x)$ garde, en effet, la forme définitive

$$m\varphi(x) + ni\varphi(1-x),$$

où m et n sont entiers positifs, nuls ou négatifs. En effet, une fois cette forme acquise, elle ne peut plus changer, puisque $i\varphi(1-x)$ varie de $\pm 2\varphi(x)$ et $\varphi(x)$ de $\pm 2i\varphi(1-x)$ à la traversée de la coupure qui lui correspond.

5. Considérons maintenant la fonction multiforme $\nu(x)$, qui, dans le *domaine initial* limité aux deux coupures $1 —— \infty$ et $0 —— -\infty$, est définie par le quotient

$$\nu(x) = \frac{i\varphi(1-x)}{\varphi(x)}.$$

Comme $\varphi(x)$ ne s'annule qu'au point critique $x = \infty$, cette fonction $\nu(x)$ n'admet que les trois points critiques des fonctions $\varphi(x)$ et $\varphi(1-x)$, à savoir 0, 1 et ∞. De plus, la fonction $\nu(x)$ est uniforme dans le domaine initial borné par ces deux coupures. Faisons maintenant varier x dans un domaine quelconque. Après un nombre quelconque de traversées de coupures, l'expression de la fonction prendra la forme

$$\frac{m'\varphi(x) + n'i\varphi(1-x)}{m\varphi(x) + ni\varphi(1-x)},$$

où les coefficients m, n, m', n' sont des entiers, liés par la relation

$$mn' - nm' = 1.$$

En effet, cette relation se vérifie dans l'expression initiale de $\nu(x)$ où $m' = n = 0$ et $m = n' = 1$. Or, quand cette relation a lieu, elle n'est pas altérée à la traversée de la coupure $1 —— \infty$ (n et n' variant alors respectivement de $\pm 2m$ et $\pm 2m'$), ni à la traversée de la coupure $0 —— -\infty$ (m et m' variant respectivement de $\pm 2n$ et de $\pm 2n'$). Donc la relation est générale.

6. La propriété la plus remarquable de la fonction $\nu(x)$ est d'avoir *sa partie imaginaire toujours positive*, sauf au point $x = \infty$ où elle est nulle. On le vérifie d'abord pour le domaine initial, de la manière suivante: si

l'on considère l'intégrale qui définit $\varphi(x)$, on voit que le facteur $1/\sqrt{u - x}$ et l'élement de l'intégrale avec lui, ont leur argument de même signe pour toutes les valeurs de u de 1 à ∞. Cet argument atteint sa plus grande valeur absolue pour $u = 1$. Donc, les arguments étant d'ailleurs supposés compris entre $\pm\,\pi$, il vient

$$\arg.\ \varphi(x) = -\frac{\theta}{2}\arg.\ (1 - x) \qquad (0 < \theta < 1),$$

On trouve, de même,

$$\arg.\ \varphi(1 - x) = -\frac{\theta'}{2}\arg.\ x \qquad (0 < \theta' < 1).$$

Mais les deux arguments précédents sont de signes contraires; leur différence qui est l'argument du quotient $\varphi(1 - x)/\varphi(x)$, sera donc en valeur absolue inférieur à la moitié de l'angle des deux vecteurs x et $1 - x$. Ces deux vecteurs, inclinés en sens contraires sur l'axe des x, font entre eux un angle $< \pi$. Donc l'argument de $\varphi(1 - x)/\varphi(x)$ est $< \pi/2$ en valeur absolue. Dans ce cas, la partie réelle de ce quotient, c'est-à-dire la partie imaginaire de $\nu(x)$, est positive. Il ne peut y avoir exception que si $\nu(x)$ s'annule, c'est-à-dire si $x = \infty$.

Cette propriété qui a lieu dans le domaine initial, subsiste dans les autres domaines. Soit, en effet,

$$\nu(x) = \alpha + \beta i, \text{ dans le domaine initial,}$$

$$\nu(x) = A + Bi, \text{ dans un autre domaine.}$$

D'après la formule du n° 5, on a, dans un domaine quelconque,

$$A + Bi = \frac{m' + n'(\alpha + \beta i)}{m + n(\alpha + \beta i)};$$

d'où

$$B = \frac{(mn' - m'n)\beta}{(m + n\alpha)^2 + n^2\beta^2} = \frac{\beta}{(m + n\alpha)^2 + n^2\beta^2}.$$

Donc B est encore positif comme β.

7. Cette dernière expression de B met en évidence un fait essentiel: *supposons que x varie de manière que β tende vers l'infini, alors la partie imaginaire B de $\nu(x)$ doit être infiniment voisine de l'une des deux limites* 0 *ou l'infini.* En effet, de quelque manière que varient m et n, deux cas seulement sont possibles: si n n'est pas nul, B est infiniment voisin de 0, cas il est de module $\leqq 1/\beta$; si n est nul, m est égal à $\pm$ 1 (à cause de la condition $mn' - m'n = 1$), donc B est égal à β et tend vers l'infini avec lui.

8. Nous allons maintenant vérifier que, *si x tend vers* 0, *β tend vers l'infini.* Comme $\nu(0) = i\varphi(1)/\varphi(0)$ et que $\varphi(0)$ est fini et positif, il faut montrer que la partie réelle de $\varphi(x)$ tend vers l'infini quand x tend vers 1. Il suffit, pour

cela, de chercher la valeur principale de l'intégrale

$$\varphi(1 - x) = \int_1^\infty \frac{du}{\sqrt{u(u-1)(u-1+x)}},$$

quand x tend vers 0. Pour obtenir cette valeur, bornons l'intégration à l'intervalle $(1, 1 + \epsilon)$ où l'élément devient infini; remplaçons le facteur u par sa valeur limite 1 pour ϵ infiniment petit. La valeur principale cherchée sera la valeur principale, pour x tendant vers 0, de l'intégrale élémentaire

$$\int_1^{1+\epsilon} \frac{du}{\sqrt{(u-1)(u-1+x)}} = \int_0^\epsilon \frac{du}{\sqrt{u(u+x)}} = \int_0^{\sqrt{\epsilon}} \frac{2du}{\sqrt{u^2+x}},$$

c'est-à-dire $-$ Log x, les termes négligés restant finis. La partie réelle $-$ Log $|x|$ est bien infinie.

9. Si l'on rapproche les conclusions des deux derniers n[os] 7 et 8, on en conclut que, *si B ne peut tendre ni vers* 0 *ni vers l'infini, x ne peut pas tendre vers* 0.

Nous avons ainsi obtenu toutes les propriétés de la fonction modulaire $\nu(x)$ qui sont nécessaires pour prouver le théorème de M. Montel. Arrivons donc à celui-ci.

10. *Définition.* Une famille F de fonctions $f(z)$ est *normale* dans un domaine D, si de toute suite infinie $f_1, f_2, \cdots$ appartenant à la famille, on peut extraire une suite convergeant uniformément vers une fonction φ ou vers l'infini, c'est-à-dire que $1/f$ tend uniformément vers 0, dans le domaine D.*

11. *Lemme. Si les fonctions d'une suite normale ne prennent pas la valeur a (qui peut être ∞), la suite ne peut converger vers a en un point de l'interieur du domaine D, sans converger vers a dans le domaine D tout entier, ce domaine étant supposé connexe.*

Comme on peut remplaçer f par $f - a$ ou par $1/f$, on peut supposer dans la démonstration $a = 0$. Si la limite φ de $f_1, f_2, \cdots, f_n, \cdots$ n'est pas nulle partout, les zéros de φ sont isolés. Soit α l'un d'eux (non situé par hypothèse sur la frontière de D). On peut, dans D, tracer un cercle γ autour de α sur lequel φ n'est pas nul, et alors on a

$$\frac{1}{2\pi i}\int_\gamma \frac{\varphi'(z)}{\varphi(z)}\,dz \neq 0,$$

car cette intégrale donne le nombre des zéros de φ dans le cercle γ. Mais ceci, d'autre part, est impossible car cette intégrale est la limite de celle-ci:

$$\frac{1}{2\pi i}\int_\gamma \frac{f_n'(z)}{f_n(z)}\,dz$$

qui est nulle en vertu du même principe.

* M. Montel dit, au lieu de cela, dans tout domaine *intérieur* à D. La modification ici introduite simplifie le discours.

Le lemme précédent, déjà établi par M. Montel, a été démontré autrement par cet auteur. Voici encore deux autres principes dûs à M. Montel. Nous en passerons les démonstrations. Celle du premier principe est facile à retrouver et le second principe se ramène au premier.

12. *Principes de M. Montel.** 1° Une famille de fonctions holomorphes, bornées dans leur ensemble dans un domaine D est normale dans tout domaine intérieur à D (au sens étroit).

2° Une famille de fonctions holomorphes dont les valeurs ne peuvent tendre vers une constante assignée a ni être égale à cette constante, dans un domaine D, est une famille normale dans tout domaine intérieur à D (au sens étroit).

13. *Théorème fondamental de M. Montel.*† *Soit F une famille de fonctions $f(z)$, holomorphes dans un domaine connexe D, limité par un ou plusieurs contours. Si les fonctions $f(z)$ ne prennent, dans D, ni la valeur* 0 *ni la valeur* 1, *je dis que la famille F est normale dans tout domaine D' interieur à D (au sens étroit).*

Observons d'abord que la démonstration se ramène au cas où D' est limité par un seul contour. En effet, par des coupures, on peut partager D' en morceaux $D_1, D_2, \cdots$ de cette nature. Or, si la famille F est normale dans chaque morceau, elle le reste dans le domaine entier $D_1 + D_2 + \cdots$ C'est presque immédiat. On peut extraire de F une suite F_1 uniformément convergente dans D_1, de celle-ci une suite F_2 convergente dans D_2 (donc dans $D_1 + D_2$) et ainsi de suite. La dernière suite formée convergera dans D'.

Considérons donc un domaine D' intérieur à D et limité par un seul contour. Formons la famille des fonctions $\nu(f)$ de z. Ces fonctions n'ont aucun point critique dans un domaine D'' à contour unique contenu dans D et contenant D', donc elles sont holomorphes dans D''. Comme la partie imaginaire de $\nu(f)$ est positive, la famille est normale, en vertu du principe 2° de M. Montel (n° 12), dans un domaine D''' contenu dans D'' et contenant encore D'.

Nous allons montrer que la famille F l'est aussi. Extrayons de cette famille une suite infinie de fonctions $f_1, f_2, \cdots, f_n, \cdots$ et considérons leurs valeurs en un point particulier z_0 du domaine D'. La suite

$$f_1(z_0),\ f_2(z_0),\ \cdots,\ f_n(z_0),\ \cdots$$

admet au moins une limite α, et l'on peut extraire de la précédente une nouvelle suite

$$f_1'(z_0),\ f_2'(z_0),\ \cdots,\ f_n'(z_0),\ \cdots$$

* Il y a un changement dans l'énoncé provenant de la modification apportée à la définition des familles normales.

† Voir la Note précédente.

n'admettant que la seule limite α et, par conséquent, convergente au point z_0.

Supposons d'abord que l'on puisse choisir α différent des trois nombres 0, 1 et ∞. Alors nous formons la suite

$$\nu(f_1'),\ \nu(f_2'),\ \cdots,\ \nu(f_n'),\ \cdots$$

Celle-ci est normale dans D''', donc on peut en extraire une suite, convergente dans D''',

$$\nu(f_1''),\ \nu(f_2''),\ \cdots,\ \nu(f_n''),\ \cdots$$

Ces fonctions ont leur partie imaginaire positive et non nulle dans D puisque f'' n'y devient pas infinie. Il s'ensuit que la partie imaginaire de ces fonctions ne peut pas converger vers 0 dans D, car alors elle devrait converger vers 0 partout en vertu du lemme du n° 11. Or cela n'a pas lieu, par hypothèse, au point z_0. De même, la partie imaginaire ne peut pas converger vers ∞, car alors (f'') devrait converger vers l'infini partout (n° 11), ce qui encore une fois n'a pas lieu au point z_0. Dans ces conditions, d'après l'observation du n° 9, f_n'' ne peut pas tendre vers 0 dans D''' et, par conséquent, la suite

$$f_1'',\ f_2'',\ \cdots,\ f_n'',\ \cdots$$

est normale dans le domaine D' qui est intérieur à D'''. La famille primitive F l'est donc aussi.

Supposons, en second lieu, que la suite $f_1(z_0),\ f_2(z_0),\ \cdots$ n'ait pas d'autre limite que l'un des nombres 0, 1, ∞. Ce cas se ramène au précédent. D'abord on peut toujours admettre que la suite tend vers 1, car, si f tend vers 0 ou vers ∞, on peut lui substituer la fonction $1 - f$ ou $1 - 1/f$ qui tend vers 1 et qui ne prend pas non plus les valeurs 0 et 1. Supposons donc que la suite $f_1(z_0),\ f_2(z_0),\ \cdots$ converge vers 1. Nous pouvons déterminer les radicaux de manière que

$$\sqrt{f_1(z_0)},\ \sqrt{f_2(z_0)},\ \cdots$$

ait pour limite 1. Alors la suite des fonctions

$$\sqrt{f_1(z)},\ \sqrt{f_2(z)},\ \cdots$$

prenant les valeurs précédentes au point z_0 est normale. En effet, les fonctions sont holomorphes, parce que $f(z)$ ne s'annule pas, et l'on peut appliquer la démonstration précédente. Donc la famille initiale F est normale aussi.

FUNCTIONS WHICH MAP THE INTERIOR OF THE UNIT CIRCLE UPON SIMPLE REGIONS.

By J. W. Alexander, II.

§ 1.

A necessary and sufficient condition that a function $w = w(z)$ which is analytic at every point within the circle $|z| = 1$ shall map the interior of the circle upon a simple region in the w plane is that

$$\text{(I)} \qquad \frac{w(z_1) - w(z_2)}{z_1 - z_2} \neq 0$$

as long as z_1 and z_2 are points within the circle. It is proposed to derive a number of other conditions, sufficient though not necessary, which can often be applied more readily than (I) in actually testing a given function. By means of these, it will be possible to determine certain classes of functions which map the interior of the unit circle upon simple regions such as, for example, the functions (§ 8)

$$\text{(1)} \qquad w = a_1 z + a_2 \frac{z^2}{2} + a_3 \frac{z^3}{3} + \cdots$$

and

$$\text{(2)} \qquad w = a_1 z + \frac{a_3 z^3}{3} + \frac{a_5 z^5}{5} + \cdots$$

where, in both cases, the a's form a non-increasing sequence of positive numbers.

In the earlier part of the paper, the case where the function $w(z)$ is a polynomial will be considered. A study will be made of the relation between the mapping of the unit circle and the position of the roots of w and of dw/dz.

§ 2.

A polynomial $w(z)$ of degree n maps the z plane conformally upon an n-sheeted Riemann surface W, except that at the branch points of W, defined by the vanishing of dw/dz, angles fail to be preserved. The unit circle in the z plane is mapped upon a curve C whose shape is determined except perhaps for a translation, rotation, and magnification when the position of the roots of $w(z)$ or of dw/dz is known. If we write

$$\text{(3)} \qquad w = cz^k(z - z_{k+1})(z - z_{k+2}) \cdots (z - z_n)$$

where $z_{k+1}, z_{k+2}, \cdots, z_n$ denote the roots of w other than 0, it becomes evident that the change in the argument of w as z varies from z' to z'' is equal to the sum of the angles generated at the roots of w by the vectors joining the latter to z, each root being counted with its proper multiplicity.*

The ratio between the length of an element of arc in the w plane and the corresponding element in the z plane, or the *distortion* of the map, is $|dw/dz|$; the angle between the two elements, or the *twist*, is arg dw/dz. If we write

$$\frac{dw}{dz} = az^{k-1}(z - \zeta_k)(z - \zeta_{k+1}) \cdots (z - \zeta_{n-1}) \tag{4}$$

we may easily visualize the distortion and twist at any point, as well as the way in which the twist varies when z moves from z' to z''.

As an immediate consequence of the above, it is clear that if all the roots of the polynomial w are on the same side of a straight line, the line is transformed into a curve which turns continuously about the origin, while if all the roots of the derived function are on the same side of a line, the line is transformed into a curve whose curvature is either everywhere positive or everywhere negative.

We shall bring these preliminaries to a close by recalling that the roots of w have the same arithmetic mean as the roots of dw/dz, for if

$$w = a_0 + \cdots + a_{n-1}z^{n-1} + a_nz^n,$$

both arithmetic means fall at $-\dfrac{a_{n-1}}{na_n}$.

§ 3.

Omitting the trivial case where the degree n of the polynomial w is unity, let us determine in the z plane a region R_n such that every polynomial w whose roots are all within or on the boundary of R_n shall map the unit circle upon a simple region. By observing for what values of a the function

$$w = (a - z)^n \tag{5}$$

fails to have this property, we can immediately set bounds to the region R_n. Thus, points within the unit circle are excluded at once, since dw/dz also vanishes at a. Moreover, since the function (5) transforms circles about a into circles about the origin and multiplies by n the angle subtended by an arc of the first of these circles at its center, the point a cannot lie so close

* This geometrical method of interpretation may, of course, be extended to the case of a function defined by a convergent infinite product

$$w = cz^k(1 - z/z_{k+1})^{l_1}(1 - z/z_{k+2})^{l_2} \cdots, \tag{3'}$$

where the l's need only be real. The angle generated at z_{k+2}, must then be multiplied by l_i.

to the unit circle that the two tangent rays from a enclose an angle of more than $2\pi/n$. Therefore, *the region R_n contains no point within the circle K of radius $1/\sin(\pi/n)$ about the origin.*

Now if we draw any tangent T to the unit circle as, for example, the vertical tangent through the point 1, we may take for the region R_n that portion of the half-plane to the right of T which lies without the circle K. For when the roots of the polynomial w are all *within* the region R_n as thus defined, the roots of dw/dz are all to the right of the tangent T, by the well-known theorem* that the roots of the derivative of a polynomial are all contained within every convex polygon which encloses all the roots of the polynomial itself. The left half-plane bounded by T is therefore mapped upon a region containing none of the branch points of W and bounded by the image of T, a curve which winds about the origin continually in the same sense. As a consequence of this, two points z' and z'' of the left half-plane correspond to the same point w only if the linear segment $z'z''$ is mapped upon a curve which encloses the origin. But this is impossible when z' and z'' are within or on the unit circle, for they then subtend an angle of less than $2\pi/n$ at each of the roots of w, or in all an angle less than 2π. If the roots of w are on the boundary of R_n, the image of the unit circle may touch itself, but the interior will still be mapped upon a simple region. The region R_n therefore has the required property.

We shall now prove that the region R_n cannot be enlarged by the addition of other points of the plane. Insofar as the points within the circle K of radius $1/\sin(\pi/n)$ are concerned, this has already been proved; there remain to be considered the points outside or on this circle and to the left of the tangent T. Every point P of the latter class may be joined to a point Q within the unit circle so near to the point 1 that when PQ is produced beyond Q to a point Q' such that $\overline{PQ'} = n\overline{PQ}$, Q' shall lie within R_n. If we then consider the polynomial w which vanishes $n - 1$ times at Q' and once at P, we find that since dw/dz vanishes $n - 2$ times at Q' it must vanish the remaining time at Q within the unit circle, by the theorem that the roots of $w = 0$ and of $dw/dz = 0$ have the same arithmetic mean. The point P therefore cannot be added to the region R_n, and we have

Condition (II). *If the roots of a polynomial $w(z)$ of the nth degree all lie within or on the boundary of a region R_n consisting of the points outside a circle of radius $1/\sin(\pi/n)$ about the origin and on the side of a tangent to the unit circle away from the circle, the interior of the circle is mapped upon a simple region. Moreover, the test no longer holds if the region R_n be enlarged.*

A function

$$w = (z - a_1)^{\alpha_1}(z - a_2)^{\alpha_2} \cdots (z - a_k)^{\alpha_k}, \tag{6}$$
$$\alpha_i > 0, \qquad i = 1, 2, \cdots, k, \qquad \Sigma\alpha_i = n$$

* A simple proof of this theorem is given by F. Irwin, these *Annals*, vol. 16 (1915), p. 138.

may be regarded as a *generalized polynomial* of degree n, where n need not be an integer. It will then be found that Condition (II) holds true for generalized as well as ordinary polynomials. In fact, the entire argument, including the proof of the theorem referred to in the footnote on p. 14, may be applied with only trivial modifications to these more general functions. The argument may also be extended to many functions defined by infinite products.

§ 4.

A region every point of which may be joined to a point a by means of a linear segment consisting only of points of the region will be called a *star-shaped region* with center at a. When an arbitrary point of the region may be chosen as center, the region will be called *convex*.

Now, let $w = w(z)$ map the unit circle upon a star-shaped region. Then arg w is a never decreasing function of $\theta = \arg z$ as z describes the unit circle in the positive sense. If $w = w(z)$ maps the unit circle upon a convex region, the slope, $\arg z(dw/dz)$, of the image of a radius of the circle is a never decreasing function of θ as z describes the unit circle, or as z describes the circle $|z| < 1 - \epsilon$ when $\arg z(dw/dz)$ fails to exist on the unit circle. We can therefore state the

THEOREM. *A necessary and sufficient condition that the function $w = w(z)$ map the interior of unit circle upon a convex region is that $z(dw/dz)$ shall map the interior of the circle upon a star-shaped region with center at the origin.*

Let us now determine the largest region S_n such that if the $n - 1$ roots of the polynomial

$$w = zP_{n-1}(z) \tag{7}$$

other than zero lie within or upon the boundary of S_n, the image of the interior of the unit circle shall be a star-shaped region.

This problem may be solved at once. Clearly, the region S_n contains no point a within a circle of radius n about the origin, for the derivative of the function

$$w = z(z - a)^{n-1} \tag{8}$$

vanishes $n - 2$ times at a and once at a/n so that we must have $|a| \geqq n$. On the other hand, if the remaining $n - 1$ roots of w are all without a circle of radius n, the angle subtended at each by an element of arc ds on the unit circle is less than $ds/(n - 1)$, therefore the sum of the angles subtended at them all is less than ds, the angle subtended at the root $z = 0$. The unit circle is therefore mapped upon a curve which turns about the origin continually in the same sense, and its interior is therefore mapped upon a star-shaped region. We therefore have the

THEOREM (CONDITION III). *The function* $w = zP(z)$, *where* $P(z)$ *is a polynomial of the* $(n-1)$*st degree maps the interior of the unit circle upon a star-shaped region with center at the origin if the roots of* $P(z)$ *all lie without or upon a circle of radius* n *about the origin. Moreover, the region* S_n *composed of the points without or on the circle cannot be enlarged by the addition of other points of the plane.*

Regarding $zP(z)$ as the function $z(dw/dz)$ in the first theorem of this section, we have the

COROLLARY. *The polynomial* $w = w(z)$ *of the* n*th degree maps the interior of the unit circle upon a convex region if the roots of* $(dw/dz) = 0$ *all lie without or upon a circle of radius* n *about the origin. Moreover, the region thus defined cannot be enlarged.*

We also have the

COROLLARY. *The polynomial* $w = w(z)$ *of the* n*th degree maps the interior of the unit circle upon a convex region if its roots all lie within or on the boundary of the half-plane which is bounded by a line at a distance* n *from the origin and which does not contain the origin.*

For the roots of $(dw/dz) = 0$ then all lie without the circle of radius n about the origin.

COROLLARY. *Let* z_1 *be a simple root of the polynomial* $w(z)$ *of the* n*th degree and let* d *be the distance from* z_1 *to the nearest other root of* $w(z)$. *Then* dw/dz *cannot vanish within a circle of radius* d/n *about* z_1. *Moreover, there exists a polynomial satisfying the conditions of the corollary and such that* dw/dz *vanishes on the circle of radius* d/n.

For, by a linear transformation, z_1 may be transformed into the origin and the circle of radius d/n into the unit circle.

The extension of the results of this section to the case of generalized polynomials suggests itself at once and need not be insisted upon.

§ 5.

The position of the roots of dw/dz also determines whether or not the interior of the unit circle is mapped by $w = w(z)$ upon a non-overlapping region, hence it ought to be possible to find the greatest region R_n' such that when the roots of dw/dz all lie within R_n', the function $w(z)$ satisfies Condition (I). This problem appears, however, to be much more difficult than the analogous one considered in § 3. We shall state a few theorems about the region R_n' without defining it completely.

The region R_n' *contains no point within a circle of radius* $1/\sin(\pi/n)$, as may be seen at once by going back to Function (5). On the other hand, a circle about the origin can easily be found outside of which there are only points of R_n'. For when all the roots of dw/dz lie without the unit circle,

the image C of the latter can overlap only if, when regarded as a curve of the w plane, it contains a closed loop winding in the negative sense about one or more branch points of the Riemann surface W. The presence of this loop is evident if we consider the variation in the image of a circle K concentric with the unit circle as the radius of K increases from 0 to 1. A double point can appear on the image of K only if the curve passes through a position in which it has a cusp or if it makes one of the negative loops in question. Moreover, a negative loop having once appeared cannot disappear unless the curve goes through a position in which it has a cusp. Now the curve cannot have a cusp at any stage since dw/dz does not vanish within the unit circle, therefore, if the curve crosses itself, it must include a negative loop. But the difference between the distortion at the beginning of the loop and that at the end must then be greater than π plus the angle subtended by the arc of the unit circle corresponding to the loop. The curve C therefore cannot cross itself if

$$\text{(IV)}\qquad \begin{gathered}\frac{dw}{dz} \neq 0 \text{ within the unit circle and} \\ \arg\left(\frac{dw}{dz}\right)_1 - \arg\left(\frac{dw}{dz}\right)_2 \leqq \pi + \arg(z_2 - z_1)\end{gathered}$$

where z_1z_2 is a positive arc on the unit circle. The second inequality may be written

$$\text{(IV}'\text{)}\qquad \arg\frac{\left(z\dfrac{dw}{dz}\right)_1}{\left(z\dfrac{dw}{dz}\right)_2} \leqq \pi.$$

A fortiori, the curve C does not cross itself if

$$\text{(V)}\qquad \arg\frac{\left(\dfrac{dw}{dz}\right)_1}{\left(\dfrac{dw}{dz}\right)_2} \leqq \pi.$$

Therefore, we have

CONDITION (VI). *If the sum of the angles subtended at the roots of dw/dz by the tangents from these points to the unit circle is less than or equal to π, the interior of the unit circle is mapped by the function $w = w(z)$ upon a simple region.*

It also follows that *the region R_n' contains all the points without or upon the circle of radius $1/\sin(\pi/2(n-1))$ about the origin.*

A better approximation of the region R_n' could of course be obtained by using Condition (IV) rather than (V), but (IV) does not appear to give

the entire region either. If we go back to the discussion in § 2, we see that the twist at the point of maximum twist on the unit circle tends to increase as the roots of dw/dz approach the unit circle and that it tends to be greater when the roots are grouped together at a point rather than evenly distributed around the circle. It is therefore highly probable that the region R_n' is determined by Function (5) and consists of the points outside the circle of radius $1/\sin(\pi/n)$ about the origin as center.

§ 6.

Condition (V) implies that the range of variation of dw/dz on the unit circle is less than π, or in other words that the function $w' = (dw(z)/dz)$ maps the interior of the circle upon a region contained within a half-plane bounded by a straight line through the origin. The problem of finding functions w' which satisfy Condition (V) therefore reduces to one already examined by Carathéodory. If w' be expanded about the origin in a series

$$w' = a_1 + a_2 z + a_3 z^2 + \cdots, \tag{9}$$

a necessary and sufficient condition that w' map the interior of the unit circle upon a region contained in a given half plane is that the real and imaginary parts of the first n coefficients $a_1, a_2, a_3, \cdots$ shall for every n be the homogeneous coordinates of a point in a certain convex domain in $2n$ space.* The a's may be made to appear in the series defining w by writing

$$w = A + a_1 z + \frac{a_2 z^2}{2} + \cdots + \frac{a_n z^n}{n} + \cdots. \tag{10}$$

As a result of Condition V, we have the

THEOREM. *If in the series (10),* $|a_1| \geqq |a_2| + |a_3| + \cdots + |a_n| + \cdots$, *the interior of the unit circle is mapped upon a simple region.*

For dw/dz always lies in the half-plane bounded by the normal to the vector a_1 through the origin and containing the point a_1.

In deriving Condition (V), we really assumed that w was a polynomial, but there is no difficulty in showing that (V) holds equally well for a convergent infinite series. That no two points within the unit circle can be mapped upon the same point w will then be seen by enclosing them within a circle interior to and concentric with the unit circle. The image of the interior of the latter circle can be approximated as closely as we please by a series of a finite number of terms.

* C. Carathéodory, "Über den Variabilitätsbereich der Koeffizienten, etc.," Math. Ann., vol. LXIV (1907), p. 95.

§ 7.

Although no root of dw/dz can lie within the unit circle if the curve C is to bound a simple region, it is possible for all the roots to lie upon the circle itself provided they are not too close to one another. Let us suppose that they are all upon the unit circle and follow one another in the order $\zeta_1, \zeta_2, \cdots, \zeta_{n-1}$ as the circle is described in the positive sense. Then every arc $\zeta_i\zeta_{i+1}$ is mapped upon a curve of positive curvature, since a point in describing the arc turns in a positive sense about all the points $\zeta_1, \zeta_2, \cdots, \zeta_{n-1}$. The curve C therefore consists of $n - 1$ arcs all concave inward and abutting in pairs at the branch points of the Riemann surface W at each of which they form a cusp of C. It may further be seen that the curve C cannot cross itself if the angle made by the tangent to C increases by at least π as we travel along the curve from one cusp to the next. Because, under these conditions, the curve C would only cross itself by containing a simple loop winding in the positive sense about the region which it enclosed. But if the curve C contained such a loop, so also would the variable curve C' defined as the image of a circle concentric with the unit circle but of radius less than 1. For as we varied the radius from 1 to 0, the loop could only disappear by shrinking to a cusp, which is impossible. On the other hand, we know that when the radius is sufficiently small, the curve C' differs by as little as we please from a circle.

Now the angle through which the tangent to C turns is equal to the angle θ subtended by the corresponding arc of the unit circle at the center of the circle plus the sum of the angles subtended at the points $\zeta_1, \zeta_2, \cdots, \zeta_{n-1}$ or

$$\theta + \frac{n-1}{2}\theta = \frac{n+1}{2}\theta.$$

The curve C therefore cannot cross itself if

$$\text{(VII)} \qquad \theta \geqq \frac{2\pi}{n+1}.$$

By applying this test, we see that the functions

$$\text{(11)} \qquad w = z - \frac{z^n}{n},$$

$$\text{(12)} \qquad w = z + \frac{z^2}{2} + \frac{z^3}{3} + \cdots + \frac{z^n}{n},$$

$$\text{(13)} \qquad w = z + \frac{z^3}{3} + \frac{z^5}{5} + \cdots + \frac{z^{2k+1}}{2k+1},$$

map the interior of the unit circle upon simple regions since their derivatives

are

$$\frac{dw}{dz} = 1 - z^{n-1}, \qquad \theta = \frac{2\pi}{n-1}, \tag{11'}$$

$$\frac{dw}{dz} = 1 + z + \cdots + z^{n-1}, \qquad \theta = \frac{2\pi}{n}, \tag{12'}$$

$$\frac{dw}{dz} = 1 + z^2 + \cdots + z^{2k}, \qquad \theta = \frac{2\pi}{n+1} \quad (n = 2k+1); \tag{13'}$$

respectively. We shall have occasion to consider the above functions presently.

§ 8.

We shall mention one more test leading to a number of applications. The interior of the unit circle will be mapped upon a simple region if it is possible to draw from every point $w(e^{i\theta})$ of the curve C a ray which satisfies the following three conditions:

(*a*) The angle through which it must be turned in order to make its direction coincide with the positive direction along the tangent to the curve C shall be between 0 and π inclusive.

(*b*) The angle ψ which it makes with the positive axis of reals shall be a never-decreasing function of θ (= arg z).

(*c*) The angle ψ shall not increase by more than 2π as θ varies from 0 to 2π.

Condition (*a*) may be expressed thus:

$$\pi \geqq \arg\left(z\frac{dw}{dz}\right) + \frac{\pi}{z} - \psi \geqq 0$$

or

$$\frac{\pi}{2} \geqq \arg\left(z\frac{dw}{dz}\right) - \psi \geqq -\frac{\pi}{2}. \tag{14}$$

It means geometrically that the ray never points to the side of C corresponding to the interior of the unit circle.

To prove that these three conditions are sufficient, we observe first that dw/dz cannot vanish within the unit circle, otherwise the tangent to the curve C would turn through an angle $2k\pi$, $(k > 1)$ as θ varied from 0 to 2π. It would then be impossible to find a set of rays satisfying the first and last conditions simultaneously. On the other hand, when the roots of dw/dz all lie without or on the circle, the curve C cannot cross itself unless it contains a closed loop winding in the negative sense about the region which it encloses (cf. § 5). The slope of the tangent then decreases by at least π as the loop is described and Conditions (*a*) and (*b*) cannot both be satisfied.

Let us apply this test to the function

$$(12'') \qquad w_n = z + \frac{z^2}{2} + \cdots + \frac{z^n}{n} \cdots$$

previously considered. If we put

$$\psi = \arg \frac{z}{1-z},$$

conditions (*b*) and (*c*) are certainly satisfied. Moreover,

$$\arg z \frac{dw_n}{dz} - \arg \frac{z}{1-z} = \arg (1 - z^n),$$

which satisfies (14) and therefore Condition (*a*). We are enabled, however, to go much further than before if we observe that ψ is independent of n, the degree of the polynomial w_n. For if two functions w_i and w_j satisfy (14) for the same choice of ψ, so also does their sum, since $\arg (d/dz)(w_i+w_j)$ lies somewhere between $\arg (dw_i/dz)$ and $\arg (dw_j/dz)$. More generally, if $c_1, c_2, \cdots, c_n$ be a set of positive constants and $w_1, w_2, \cdots, w_n$ a set of functions satisfying (14) for the same choice of ψ, the sum function

$$(15) \qquad w = c_1 w_1 + c_2 w_2 + \cdots + c_n w_n$$

also satisfies (14).

THEOREM. *The function*

$$w = a_1 z + \frac{a_2 z^2}{2} + \cdots + \frac{a_n z^n}{n},$$

where the a's form a non-increasing sequence of positive numbers maps the interior of the unit circle upon a simple region.

For w may be written as

$$w = a_n w_n + (a_{n-1} - a_n) w_{n-1} + \cdots + (a_1 - a_2) w_1$$

where w_n is the function defined in (12″). The function w is thus seen to be of the form (15).

Starting from

$$(13) \qquad w = z + \frac{z^3}{3} + \cdots + \frac{z^n}{n}, \qquad n = 2k + 1$$

and putting $\psi = \dfrac{1+z}{1-z}$, we have

$$\arg z \frac{dw}{dz} - \arg \frac{1+z}{1-z} = \arg z \frac{(1-z^n)}{(1+z)^2} = \arg (1 - z^n)$$

which also satisfies (14). This leads us to the

THEOREM. *The function*

$$w = a_1 z + \frac{a_3 z^3}{3} + \cdots + \frac{a_{2k+1} z^{2k+1}}{2k+1},$$

where the a's form a non-increasing sequence of positive numbers maps the interior of the unit circle upon a simple region.

Other analogous theorems will easily suggest themselves. By the method indicated at the end of § 6, we may pass to the case of infinite series and obtain the theorems mentioned in the introduction.

§ 9.

By the method of § 8 and starting from the function (11), we could have proved the theorem of § 6. We shall now prove the stronger

THEOREM. *The function*

$$w = a_1 z + \frac{a_2 z^2}{2} + \cdots + \frac{a_n z^n}{n} + \cdots,$$

where $| a_1 | \geqq \sum_{i=2}^{\infty} | a_i |$ *maps the interior of the unit circle upon a star-shaped region with center at 0.*

To prove this theorem, it will be sufficient to show that $\arg w$ is a never decreasing function of $\theta = \arg z$. But

$$\begin{aligned} \frac{d}{d\theta} \arg w &\geqq | a_1 | \left| \frac{d}{d\theta} \arg z \right| - \sum_{i=2}^{\infty} | a_i | \left| \frac{d}{d\theta} \arg \frac{z^i}{i} \right| \\ &\geqq | a_1 | - \sum_{i=2}^{\infty} | a_i | \\ &\geqq 0, \end{aligned}$$

which proves the theorem.

As a corollary, we have the

THEOREM. *The function*

$$w = a_1 z + \frac{a_2 z^2}{2^2} + \cdots + \frac{a_n z^n}{n^2} + \cdots,$$

where $| a_1 | \geqq \sum_{i=2}^{\infty} | a_i |$ *maps the interior of the unit circle upon a convex region.*

For $z(dw/dz)$ maps the circle upon a star-shaped region with center at 0.

The author is happy to express to Professor T. H. Gronwall his indebtedness for much useful advice during the preparation of this paper.

PRINCETON, N. J.

THE ITERATION OF FUNCTIONS OF ONE VARIABLE.

By Albert A. Bennett.

A large part of the theory of the iteration of functions of several variables is illustrated by the relatively simple case of functions of one variable, and it is to the study of this case that the present paper is devoted. Several distinct classes of topics immediately suggest themselves in this connection. One of these classes comprises subjects of a purely formal nature, and is concerned only with expansions in the neighborhood of certain significant points. Another class of topics deals with the iteration of a real function within an interval. Still a third treats of the total behavior of the transforming function $B(x)$, when the given function $A_1(x)$ is analytic and defined for all values of x. Except for certain simple cases the function $B(x)$ cannot be one-valued. The present paper deals almost exclusively with the first two of the three classes of subjects just mentioned. In the formal problem, certain new points of view, and new results and formulæ are obtained, although a large proportion of this part is a systematization of results already known but largely unrelated. The second part of this paper deals with the iteration of a real function, a question which seems to have been neglected up to date.* The third topic mentioned above, has not here been touched upon, except in the case of certain simple rational functions. A few known theorems concerning the iteration of functions of a single variable, have been omitted, in view of the fact that they seem to have little significance in connection with the subject of the iteration of functions of several variables. Only those subjects have here been discussed which appear as natural introductory material to the more general case.

For references to those parts of the formal problem which have already been treated in the literature, one may consult the following:

Pincherle, on "Functional Equations and Operations," in the Encyklopädie d. Math. Wiss., II, A 11, and Encyclopédie d. Sci. Math., II, 26.

E. Schroeder, Math. Ann., 2 (1870), p. 317, and 3 (1871), p. 296.

J. Farkas, Journ. de Math. (3), 10 (1884), p. 102.

G. Koenigs, Bull. Sc. Math. (2), 7 (1883), p. 340, and (3) 1 (1884), suppl., p. 14.

* Since this paper was written, there has been a report on one phase of the case of real iteration by Mr. J. F. Ritt. Cf. Am. Math. Soc. Bull., 21 (1915), p. 379.

E. Podetti, Giorn. Mat. (2), 4 (1897), p. 264.

C. Formenti, Reale Ist. Lomb. Rend. (2), 8 (1875), p. 275.

A. N. Korkine, Bull. Sc. Math. (2), 6 (1882), p. 228.

C. Bourlet, Toulouse Ann. (1), 12 (1898), mém. no. 3, p. 1–3.

A. Grévy, Thesis, Paris, 1894, and Ann. Ec. Norm. (3), 11 (1894), p. 287.

L. Leau, Thesis, Paris, 1897.

E. M. Lemeray, C. R., 125 (1897), p. 524, also S. M. F. Bull., 26 (1898), p. 10.

O. Spiess, Math. Ann., 62 (1906), p. 226.

M. Koppe, Die Iteration des Sinus . . ., Berlin, 1909.

Pincherle and Amaldi, Le Operazione Distributive, Chap. XIV. This may be consulted in connection with part of the use here made of matrices.

Note. The matrices and iteration discussed in this paper have no relation to the matrices and "Schapira's Iteration" treated in several recent papers by L. v. David, in the J. fur Math.

PART ONE. THE ITERATION OF POWER SERIES.

Definition and Examples of Iteration.

2. Let there be given an analytic function, $A_1(x)$. The function $A_1[A_1(x)]$, we shall define as $A_2(x)$, and in general for n, a positive integer, we shall define $A_n(x)$ as identical with $A_{n-1}[A_1(x)]$, or, what amounts to the same thing, as identical with $A_1[A_{n-1}(x)]$. We may similarly denote the inverse of $A_1(x)$ by $A_{-1}(x)$. For two given functions $A_1(x)$, and $F_1(x)$ we may define a $G_1(x)$, so that $G_1(x) \equiv F_{-1}\{A_1[F_1(x)]\}$, or, in other words, so that $G_1(x)$ is the transform of $A_1(x)$ through $F_{-1}(x)$. The $G_1(x)$ and $A_1(x)$ will then be so related that for any positive integer, or even negative integer value of n, $G_n(x) \equiv F_{-1}\{A_n[F_1(x)]\}$. In particular, we may replace the consideration of $y = A_1(x)$ by that of $\eta = G_1(\xi)$, where η is obtained from y, and ξ from x, by the same linear fractional transformation.

It is natural to inquire whether we might not suppose $A_1(x)$ and $G_1(x)$, as given, and determine a transforming function F, such that $G_1(x) \equiv F_{-1}\{A_1[F_1(x)]\}$, or, as it may be written, $A_1[F_1(x)] \equiv F_1[G_1(x)]$. In particular, we may examine the case in which $G_1(x)$ is simply $x + 1$. The function F associated with A by this particular choice of G, we shall call B. We seek therefore to determine a function B, such that

$$A_1[B(x)] \equiv B(x+1). \tag{1}$$

We might equally well have started with $G_1(x)$ as cx, where c is a constant, $\neq 0$, $\neq 1$, and have sought a function $E(x)$, such that

$$A_1[E(x)] \equiv E(cx). \tag{2}$$

The two problems are not, of course, independent. We need merely replace x in (2) by c^x, to obtain

$$A_1[E(c^x)] \equiv E(c^{x+1}),$$

so that $E(c^x)$ is of the form of $B(x)$. If a function $B(x)$ be found, we may define $A_n(x)$, as identical with $B_1[n + B_{-1}(x)]$, for n not only an integer, but, indeed, for all complex values of n, and similarly for $E(x)$. The definition of $A_n(x)$, will depend to a considerable degree upon the particular function B or E, that is selected, except for n an integer. We shall define $A_n(x)$ as the nth *iterate* of $A_1(x)$, and the process of repeatedly substituting $A_1(x)$ in place of x, or, in general, of determining $A_n(x)$ from $A_1(x)$, as the process of *iteration*.

We shall now give a few simple examples of iteration, for which no extended discussion will be necessary. For the degenerate case of $A_1(x)$ equal to a constant, equation (1) defines $B(x)$ as also equal to a constant. For $A_1(x) \equiv x$, $B(x)$ may be any periodic function $P(x)$, with unity for a period. For $A_1(x) \equiv x + 1$, $B(x)$ is of the form $x + P(x)$, where $P(x)$ is defined as above. For $A_1(x) \equiv cx$, $0 \neq c \neq 1$, $B(x)$ is of the form $e^{[x+P(x)]\log c}$, $P(x)$ as before. This includes such expressions as ac^x, where a is an arbitrary nonzero constant, and more generally, it includes an extensive class of elliptic theta functions. For $A_1(x)$, an arbitrary linear fractional function, we obtain, either directly, or by reducing to a normal form, a class of associated functions, $B(x)$, containing an arbirtary periodic function $P(x)$, with unity for a period.

If $A_1(x)$ be not a linear fractional function, then either $A_1(x)$ itself, or else its inverse, $A_{-1}(x)$, must be multiple-valued, which would suggest that no one-valued function, $B(x)$, could exist for such a function $A_1(x)$. This is not, however, the case. For example, if we define $A_1(x)$ by the implicit equation:

$$\left.\begin{array}{lll} c_{00} & + c_{10}x & + c_{20}x^2 \\ + c_{01}A_1(x) & + c_{11}xA_1(x) & + c_{21}x^2A_1(x) \\ + c_{02}A_1^2(x) & + c_{12}xA_1^2(x) & + c_{22}x^2A_1^2(x) \end{array}\right\} = 0,$$

where $c_{ij} = c_{ji}$, $i, j, = 1, 2, 3$, and where the determinant of the coefficients is different from zero, we shall have a non-degenerate quadri-quadric relation between $A_1(x)$ and x. For this choice of $A_1(x)$, there exists among the solutions of (1), a $B(x)$, which is a single-valued elliptic function. The corresponding $A_n(x)$ satisfies an analogous quadri-quadric relation for every particular choice of n, whether real or complex. This is the problem which arises in connection with Poncelet Polygons, and the addition theorem of elliptic functions.*

* Cf. Cayley, Elem. Treat. on Ell. Funct., p. 340, and Halphen, Traité des Fonc. Ell., vol. II, Chaps. 9 and 10, and an article by the author in *Annals of Math.* (2), 16 (1915), p. 97–118.

The question arises as to what is the most general algebraic expression $A_1(x)$, which is defined by equating to zero a polynomial in $A_n(x)$ and x, whose coefficients vary with n, but whose degree is, in general, the same, and for which there exists a $B(x)$, satisfying (1), and one-valued all over the complex plane. This problem involves the question as to what types of algebraic curves admit of one to one continuous transformations into themselves. As is well known, the rational and the elliptic curves are the only ones admitting such transformations. The problem just proposed reduces to the study of addition theorems of rational and of elliptic functions. We shall not go further into this matter, but shall proceed immediately to a more general problem.

Significance of the Formal Iteration of Power Series.

3. If we confine ourselves to the immediate neighborhood of a finite value of x, the equation (1) imposes no restrictions on $B(x)$, and is without significance. It is only when x is in the neighborhood of infinity, that we can expect to draw formal inferences with regard to $B(x)$, as a consequence of the behavior of $A_1(x)$. We shall seek to determine a solution $B(x)$, which at least admits asymptotic representation along the positive real axis, using the term "asymptotic representation," in its widest sense. In particular, we shall require that $\lim_{x=+\infty} B(x)$ exists. By replacing $A_1(x)$, if necessary, by a linear transform of it, we may secure that $\lim_{x=+\infty} B(x) = 0$, or from (1), that $\lim_{x=0} A_1(x) = 0$, when x approaches zero along a suitable path. There is another reason, also, which induces us to require that $\lim_{x=0} A_1(x) = 0$ for at least one method of approach, a reason, which is not explicitly related to the nature of the transforming function $B(x)$. The problem of computing the coefficients of a power series representation of $A_2(x)$ from the coefficients of a series representing $A_1(x)$ requires that at least one point (e) be known, either finite or infinite, for which $A(e) = e$, unless we are to have recourse to infinite summations which cannot be convergent except when the circle of convergence for the series $A_1(x)$, includes the point corresponding to the constant term of this series.

The derivation of $B(x)$ from $A_1(x)$ involves two steps; the first, that of the purely formal derivation of the coefficients, the second, that of proving the convergence or divergence of the series formally obtained for $B(x)$. The first of these steps is not, however, entirely devoid of significance, even when the resulting series is divergent. The series

$$A_1(x) \equiv a_{11}x + a_{12}x^2 + \cdots + a_{1n}x^n + \cdots$$

may be divergent, and yet it may well happen that there exists an analytic

function $F(x)$, with an isolated essential signularity at the origin, for which

$$\lim_{x=0}\left[\frac{1}{n!}\frac{d^nF(x)}{dx^n}\right] = a_{1n},$$

along a suitably chosen path of approach. This is the case presented by the classical illustration of $y = e^{-1/x^2}$, where the path may be taken as either real semi-axis. Were we to choose infinity instead of the origin as the point approached, we should have the usual case of asymptotic convergence.* The set of coefficients $a_{11}, a_{12}, a_{13}, \cdots, a_{1n}, \cdots$, determines what may be called a *differential element of infinite order*. Two distinct functions may admit of the same asymptotic representation for a given method of approaching infinity, or to use more general ideas, they may have the same differential element at a given point, for a given path of approach. The two functions then represent curves which along a given path in the complex plane, have contact at the given point, of infinite order. So far as the formal determination of the series is concerned, we may suppose ourselves, throughout, as interested only in "asymptotic" forms, i. e., in differential elements for which questions of convergence are without significance.* It has been shown and it may be immediately verified that the analytic functions represented by

$$\sum_{n=1}^{\infty}\frac{a_n}{n!}(1 - e^{-\frac{1}{b_nx^2}})x^n,$$

where $1 < b_n > | a_n |$, b_n and a_n being real, have in general, an essential singularity at the origin, but have for real values the common differential element $(a_1, a_2, \cdots, a_n, \cdots)$, at the origin.

It would, of course, be possible to consider series $A_1(x)$, with no absolute term, but with terms containing other than merely integral powers of x. These, however, we shall not investigate. Even in the case of direct power series, a classification of types will be found necessary. We might enunciate an apparently arbitrary classification, justified *a posteriori* by the different conclusions obtained. We shall however find that the types which arise are essentially those which are suggested by regarding iteration as a case of collineation in an infinite number of dimensions. It is this point of view that we shall now present.

The Iteration of Power Series as Matrix Multiplication. Classification of Types.

4. For many purposes, it is more convenient to replace the nonhomogeneous relation $y = A_1(x)$ by the pair of homogeneous equations, obtained

* Cf. W. B. Ford, S. M. F. Bull, 39 (1911), p. 347.

† Cf. Kasner, "Conformal Geometry," Fifth Int. Cong. of Math. Proc., vol. II, p. 81 ff. The term, "differential element of infinite order," has been discussed by Kasner. See also J. F. Ritt, Bull. Am. Math. Soc., 21 (1915), p. 379.

by replacing x by x_1/x_0, and y by y_1/y_0. We shall indeed write

$$\begin{cases} y_1 = x_0 A_1\left(\dfrac{x_1}{x_0}\right), \\ y_0 = x_0. \end{cases}$$

We define $A_1^{(m)}(x)$ as the polynomial of the mth degree in x, which is obtained by neglecting all but the first m terms of $A_1(x)$. Now the first m terms of $A_n(x)$, for n a positive or negative integer, depend only upon the first m terms of $A_1(x)$, that is, $A_n^{(m)}(x)$ is determined from $A_1(x)$ by using only $A_1^{(m)}(x)$.

Every binary form of the mth degree in y_0, y_1, will be itself a binary form of the mth degree in x_0, x_1, together with an infinite series, every term of which contains x_0 to some positive power in the denominator. We shall now define $\xi_i^{(m)}$ and $\eta_i^{(m)}$ as follows:

$$\xi_i^{(m)} = x_0^i \cdot x_1^{m-i} \quad i = 0, 1, 2, \cdots, m,$$

$$\eta_i^{(m)} = \text{integral part of series for } y_0^i \cdot y_1^{m-i} \equiv x_0^i[x_0 A_1(x_1/x_0)]^{m-i}.$$

Here $\xi_i^{(m)}$ is a product transform of the variables x_0, and x_1, and $\eta_i^{(m)}$ is related to the product transform of y_0 and y_1.* The expressions $\eta_i^{(m)}$, $i = 0, 1, 2, \cdots, m$, are linear in the $m + 1$ homogeneous variables, $\xi_i^{(m)}$, $i = 0, 1, 2, \cdots, m$, and in particular $\eta_0^{(m)} = \xi_0^{(n)}$. The matrix of this linear transformation from the $(m + 1)\xi^{(m)}$'s to the $(m + 1)$, $\eta^{(m)}$'s we may represent by $\boldsymbol{A}_1^{(m)}$.

For any other series $y = H(x)$ we may make an analogous discussion, and denote by $\boldsymbol{H}^{(m)}$ the analogous matrix. The first m terms of $z = H[A_1(x)]$, where $H(0) = 0$, depend upon $\boldsymbol{A}_1^{(m)}$ and $\boldsymbol{H}^{(m)}$ alone. Let us now define $\zeta_i^{(m)}$ as the integral part of the series for $z_0^i . z_1^{m-1} \equiv y_0^i[y_0 H(y_1/y_0)]^{m-i}$, where $z_1 = y_0 H(y_1/y_0)$, and $z_0 = y_0$. We may determine the first m terms of $z = H[A_1(x)]$ in three steps, by first taking the m leading terms of $y = A_1(x)$, secondly, taking the m leading terms of $z = H(y)$, and thirdly, substituting the first m terms of $A_1(x)$ for y in the first m terms of $z = H(y)$, and taking the m leading terms of the result. This corresponds to considering, first, the symbolic equation $\eta^{(m)} = \boldsymbol{A}_1^{(m)}\xi^{(m)}$, secondly the symbolic equation $\zeta^{(m)} = \boldsymbol{H}^{(m)}\eta^{(m)}$, and finally, substituting $\boldsymbol{A}_1^{(m)}\xi^{(m)}$ for $\eta^{(m)}$ in the equation for $\zeta^{(m)}$. Since $\boldsymbol{A}_1^{(m)}$ and $\boldsymbol{H}^{(m)}$ are the matrices of linear transformations, the result of the substitution yields merely a new matrix of the same type, which is the usual matrix product, $\boldsymbol{H}^{(m)}\boldsymbol{A}_1^{(m)}$. If in particular $\boldsymbol{H}^{(m)} = \boldsymbol{A}_{n-1}^{(m)}$, then we see that $\boldsymbol{A}_n^{(m)} = \boldsymbol{A}_{n-1}^{(m)}\boldsymbol{A}_1^{(m)}$, where the product is the usual matrix product. The significance of this relation lies of course in the fact that the initial first minor of $\boldsymbol{A}_n^{(m+1)}$ is simply $\boldsymbol{A}_n^{(m)}$ itself.

* Cf. A. Hurwitz, Math. Ann., 45 (1894), p. 388, for the question of product transforms.

The relation $G_1(x) = F_{-1}\{A_1[F_1(x)]\}$ is equivalent to the equation $\boldsymbol{G}_1^{(m)} = \boldsymbol{F}_1^{(m),-1}\boldsymbol{A}_1^{(m)}\boldsymbol{F}_1^{(m)}$, where $\boldsymbol{F}_1^{(m),-1}$ is the inverse of $\boldsymbol{F}_1^{(m)}$ in the notation of matrices. This is the type of equation that arises in the theory of the classification of the linear transformations $\boldsymbol{A}_1^{(m)}$. Although the matrix $\boldsymbol{F}_1^{(m)}$ is not of the most general form possible, still it is not surprising that many of the theorems concerning the normal forms of linear transformations have here an application. For the sake of concreteness, we now give the form of $\boldsymbol{A}_1^{(5)}$

$$\boldsymbol{A}_1^{(5)} = \left\{\begin{matrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ 0 & 0 & a_{11}^2 & 2a_{11}a_{12} & 2a_{11}a_{13} + a_{12}^2 & 2a_{11}a_{14} + 2a_{12}a_{13} \\ 0 & 0 & 0 & a_{11}^3 & 3a_{11}^2a_{12} & 3a_{11}^2a_{13} + 3a_{11}a_{12}^2 \\ 0 & 0 & 0 & 0 & a_{11}^4 & 4a_{11}^3a_{12} \\ 0 & 0 & 0 & 0 & 0 & a_{11}^5 \end{matrix}\right\}.$$

The case of $\boldsymbol{A}_1^{(5)}$ is typical of $\boldsymbol{A}_n^{(m)}$ to the extent that the first row contains only zero except for a one in the initial position, the elements below the main diagonal are all zero, while the elements in the main diagonal are the successive powers of a_{n1}. The roots of the characteristic equation of the matrix $\boldsymbol{A}_1^{(m)}$ are clearly a_{11}^i, $i = 0, 1, 2, \cdots, m$. We shall now examine the normal forms for the linear transformations whose matrices are $\boldsymbol{A}_1^{(m)}$.

The first case to consider is that in which all of the roots of the characteristic equation are distinct. This is the case in which a_{11} is not an integral root of unity. We may without exception in this case* find matrices $\boldsymbol{E}^{(m)}$ by the method of undetermined coefficients, for which

$$\boldsymbol{A}_1^{(m)} = \boldsymbol{E}^{(m),-1}\tilde{\boldsymbol{A}}_1^{(m)}\boldsymbol{E}^{(m)},$$

where $\tilde{\boldsymbol{A}}_1^{(m)}$ has the same main diagonal as $\boldsymbol{A}_1^{(m)}$, but has only zeros above as well as below this diagonal. The matrix $\boldsymbol{E}^{(m)}$ has a single arbitrary constant which is determined completely, if, for example, we put $e_1 = 1$. The only invariant of $\boldsymbol{A}_1^{(m)}$ as m takes on successively greater integer values, is in this case a_{11}.

We next come to those nonsingular cases in which all of the roots of the characteristic equation coincide. This occurs when $a_{11} = 1$. We may now have as many as two algebraic and one arithmetic invariant. It is impossible to have more, however, no matter how great m be taken. When

* This first case has been treated by nearly all of the writers on the subject, Schroeder, Korkine, Farkas, Koenigs, etc., already referred to. This use of matrices $\boldsymbol{A}_n^{(m)}$ was made for the first time by H. v. Koch, Bil. t. Sv. Vet. Ak. Hand. 1, Math. 25 (1900), mém. no. 5, p. 1–24.

$a_{1i} = 0$, for every $i > 1$, we have the case of the identity. We shall have in all other cases, an arithmetic invariant k, such that $a_{1i} = 0$, $1 < i \leq k$, while $a_{1,k+1} \neq 0$. This is an invariant in the sense that no matter how $\boldsymbol{F}^{(m)}$ be chosen, the positive integer k for $\boldsymbol{G}_1^{(m)}$ where $\boldsymbol{G}_1^{(m)} = \boldsymbol{F}^{(m),-1}\boldsymbol{A}_1^{(m)}\boldsymbol{F}^{(m)}$, will be the same as for $\boldsymbol{A}_1^{(m)}$. By the use of undetermined coefficients we may find a matrix $\boldsymbol{E}^{(m)}$ such that

$$\boldsymbol{A}_1^{(m)} \equiv \boldsymbol{E}^{(m),-1}\tilde{\boldsymbol{A}}_1^{(m)}\boldsymbol{E}^{(m)},$$

where this time we mean by $\tilde{\boldsymbol{A}}_1^{(m)}$ the matrix obtained from a series of the form

$$A_1(x) \equiv x + \tilde{a}_{1,k+1}x^{k+1} + \tilde{a}_{1,2k+1}x^{2k+1} + \cdots + \tilde{a}_{1,lk+1}x^{lk+1} + \cdots,$$

and where $(1/x)\tilde{A}_1(x)$ is a power series in x^k. Both $\tilde{a}_{1,k+1}$ and $\tilde{a}_{1,2k+1}$, are invariants of $A_1(x)$ in the sense that any two series $\tilde{A}_1(x)$ equivalent to the same series $A_1(x)$, must have the same $\tilde{a}_{1,k+1}$ and $\tilde{a}_{1,2k+1}$.*

Of the nonsingular cases there remain to be considered those in which some but not all of the roots of the characteristic equation coincide. For these cases a_{11} must be an integral root of unity and we may write $a_{11}{}^h = 1$, where h is the smallest positive integer for which this is true. But if $A_1(x)$ be of this form then $A_h(x)$ will be of the type already mentioned, since it will commence with $x + \cdots$. Thus we have as invariants at least two integers h and k, and two algebraic invariants $\tilde{a}_{h,k+1}$ and $\tilde{a}_{h,2k+1}$. Here k must be a multiple of h but need not be h itself. Furthermore by the use of the method of undetermined coefficients, these are found to be the only invariants.

The only other cases which can arise are the singular cases, and in these, $a_{11} = 0$. If every succeeding a_{1i}, $i = 2, 3, \cdots$, vanishes, the expression $A_1(x)$ vanishes identically and is independent of x. Let us suppose then that $a_{1i} = 0$, $i = 1, 2, \cdots, k - 1$, while $a_{1k} \neq 0$. In this case k will be an arithmetic invariant, and this is the only invariant.† When once a normal matrix $\tilde{\boldsymbol{A}}_1^{(m)}$ has been selected, the remaining problem is to express, if possible, $\tilde{\boldsymbol{A}}_n^{(m)}$ as a matrix whose elements are explicit analytic functions of n.

Synopsis of Types.

6. On the basis of the above discussion we are prepared to enunciate‡ a classification of power series of the form

$$A_1(x) \equiv a_{11}x + a_{12}x^2 + \cdots + a_{1r}x^r + \cdots$$

from the point of view of iteration, into eleven types. We shall later examine the various cases in greater detail.

* For this and the following case, see Leau, loc. cit. Cf. also Kasner, loc. cit.

† This is the case particularly investigated by Grèvy, loc. cit.

‡ This is given here for the first time.

Case I. a_{11} is not an integral root of unity and does not vanish.

Type I*a*, $|a_{11}| < 1, \neq 0$.

Type I*b*, $|a_{11}| > 1$.

Type I*c*,* $|a_{11}| = 1$, but $a_{11}^h \neq 1$, $h = 1, 2, \cdots$.

Case II. $a_{11} = 1$.

Type II*a*, $a_{1i} = 0$, $i = 2, 3, \cdots$.

Type II*b*, $a_{1i} = 0$, $i = 2, 3, \cdots, k$, but $a_{1,k+1} \neq 0$, and $2\tilde{a}_{1,2k+1} = (k+1)\tilde{a}_{1^2,k+1}$, where $\tilde{a}_{1,k+1}$ and $\tilde{a}_{1,2k+1}$ are the invariants of the normal form already mentioned.

Type II*c*, $a_{1i} = 0$, $i = 2, 3, \cdots, k$, but $a_{1,k+1} \neq 0$, and $2\tilde{a}_{1,2k+1} \neq (k+1)\tilde{a}_{1^2,k+1}$, where $\tilde{a}_{1,k+1}$ and $\tilde{a}_{1,2k+1}$ are defined as above.

Case III. $a_{11} \neq 1$, but $a_{11}^h = 1$, where h is the smallest positive integer for which this is true.

Type III*a*, $a_{hi} = 0$, $i = 2, 3, \cdots$.

Type III*b*, $a_{hi} = 0$, $i = 2, 3, \cdots, k$, but $a_{h,k+1} \neq 0$, and $2\tilde{a}_{h,2k+1} = (k+1)\tilde{a}_{h^2,k+1}$, these being defined as above.

Type III*c*, $a_{hi} = 0$, $i = 2, 3, \cdots, k$, but $a_{h,k+1} \neq 0$, and $2\tilde{a}_{h,2k+1} \neq (k+1)\tilde{a}_{h^2,k+1}$, defined as above.

Case IV. $a_{11} = 0$.

Type IV*a*, $a_{1i} = 0$, $i = 1, 2, \cdots$.

Type IV*b*, $a_{1i} = 0$, $i = 1, 2, \cdots, k-1$, but $a_{1k} \neq 0$.

Commutative Series.

6. If the approximate structure of a series for $A_n(x)$ be once found in a more or less indefinite fashion, it is sometimes possible to obtain the coefficients of $A_n(x)$, as explicit functions of n, without first reducing $A_1(x)$ to a normal form $\tilde{A}_1(x)$. This is done by using the following considerations. If a series $A_n(x)$ exists, it must satisfy the commutative condition

$$A_n[A_1(x)] \equiv A_1[A_n(x)].$$

Whenever the solutions $C(x)$, of the equation,

$$C[A_1(x)] \equiv A_1[C(x)] \tag{3}$$

form a one-parameter family, we may seek to identify this parameter with the n of $A_n(x)$. If, however, the C's form a discrete set, or contain more than one parameter, we cannot obtain $A_n(x)$ by this means. In terms of matrices, we have the equation,

$$\boldsymbol{C}^{(m)}\boldsymbol{A}_1^{(m)} = \boldsymbol{A}_1^{(m)}\boldsymbol{C}^{(m)}$$

which obviously greatly restricts $\boldsymbol{C}^{(m)}$. It is found that except in those cases in which a_{11} is an integral root of unity, there always exists a unique

* For questions of convergence this yields three quite distinct subcases, as we shall see later.

one-parameter family of series $C(x)$, commutative with $A_1(x)$, in the above sense, provided that in the cases for which $a_{11} = 1$, we require also that $c_1 = 1$. In the cases in which $a_{11} = 0$, the series C, although existing, are not, in general, in integral powers of x.

The identification of the parameter, with n, is, of course, not unique, but involves an arbitrary periodic function of period unity. A particularly simple choice is in each case almost immediately obvious. For instance, to illustrate the most common case, let $0 \neq | a_{11} | \neq 1$, and let

$$C(x) = c_1 x + c_2 x^2 + c_3 x^3 + \cdots$$

while $A_1(x)$ is denoted as heretofore. Then $C[A_1(x)]$ will be of the form,

$$\begin{aligned} c_1 a_{11} x + c_1 a_{12} x^2 &+ c_1 a_{13} x^3 + \cdots \\ + c_2 a_{11}^2 x^2 &+ 2 c_2 a_{11} a_{12} x^3 + \cdots \\ &+ c_3 a_{11}^3 x^3 + \cdots \\ &\qquad + \cdots \end{aligned}$$

and $A_1[C(x)]$ will be obtained from this by interchanging a_{1i} and c_i, $i = 1, 2, 3, \cdots$. Equating coefficients of like powers of x, we obtain,

$$\begin{aligned} c_1 a_{11} &= a_{11} c_1, \\ c_1 a_{12} + c_2 a_{11}^2 &= a_{11} c_2 + a_{12} c_1^2, \\ c_1 a_{13} + 2 c_2 a_{11} a_{12} + c_3 a_{11}^3 &= a_{11} c_3 + 2 a_{12} c_1 c_3 + a_{13} c_1^3, \\ \cdots & \cdots \end{aligned}$$

Hence

$$c_2 = \frac{c_1}{a_{11}} \frac{(c_1 - 1)}{(a_{11} - 1)} a_{12},$$

$$c_3 = \frac{c_1}{a_{11}} \frac{(c_1^2 - 1)}{(a_{11}^2 - 1)} a_{13} + 2 \frac{c_1}{a_{11}} \frac{(c_1 - 1)}{(a_{11} - 1)} (c_1 - a_{11}) a_{12}^2.$$

$$\cdots \cdots$$

If $C(x)$ is to be identified with $A_n(x)$, we must have for integer values of n, $c_1 = a_{11}^n$. The most general form of c_1 is then, $a_{11}^{n+P(n)}$, where $P(n)$ is a periodic function of period unity, vanishing for n an integer. The most obvious choice of c_1 is undoubtedly that in which $P(n)$ vanishes identically, so that for all values of n, $c_1 = a_{11}^n$. Every succeeding c_i, $i = 2, 3, \cdots$, is then completely determined.

Discussion of the Different Types.

7. For any type under Case I, we may always determine a series $E(x)$, and if we put $e_1 = 1$, it will be determined uniquely and in such a way,

that

$$A_1(x) \equiv E_{-1}[a_{11}E(x)].$$

By writing $E_{-1}(a_{11}{}^x) \equiv B(x)$, we obtain a $B(x)$, for which

$$A_n(x) \equiv B[n + B_{-1}(x)],$$

as was desired. This gives $A_n(x)$ as an explicit function of n. We may obtain, however, another form of $A_n(x)$, which proves for many purposes much more convenient than that just given. It is to be expected that by the use of the expressions $A_1(x), A_2(x), \cdots, A_p(x), \cdots$, where p is a positive integer, we may obtain a sort of interpolation formula, giving $A_n(x)$ for all values of n. We shall seek indeed a series of the following form

$$A_n(x) \equiv \sum_{i=0}^{\infty} \alpha_i U_i(x),$$

where $U_i(x)$ is independent of n, and is a power series in x commencing with a term in x^{i+1}, expressible as a linear function of $A_0(x), A_1(x), \cdots, A_i(x)$, with coefficients which are themselves polynomials in a_{11} with numerical coefficients, and where α_i is a function of a_{11} and n alone. These conditions serve to determine the U's and α's as follows:

$$\begin{aligned} U_0(x) &= A_0(x) \equiv x, \\ U_1(x) &= A_1(x) - a_{11}A_0(x), \\ U_2(x) &= [A_2(x) - a_{11}A_1(x)] - a_{11}{}^2[A_1(x) - a_{11}A_0(x)], \\ &\cdots\cdots\cdots\cdots \\ U_i(x) &\equiv U_{i-1}[A_1(x)] - a_{11}{}^i U_{i-1}(x). \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

The coefficients, α_i, satisfy the recursion formula,

$$\alpha_i = \frac{1}{a_{11}} \frac{(a_{11}{}^{n-i+1} - 1)\alpha_{i-1}}{(a_{11}{}^i - 1)}, \quad \alpha_0 = a_{11}{}^n.$$

Then we will have

$$\begin{aligned} A_n(x) \equiv a_{11}{}^n \Big[& U_0(x) + \frac{1}{a_{11}} \frac{(a_{11}{}^n - 1)}{(a_{11} - 1)} U_1(x) \\ & + \frac{1}{a_{11}{}^2} \frac{(a_{11}{}^n - 1)}{(a_{11} - 1)} \frac{(a_{11}{}^{n-1} - 1)}{(a_{11}{}^2 - 1)} U_2(x) + \cdots \Big]. \end{aligned} \tag{4}$$

To prove that the $A_n(x)$ determined by (4) is actually identical with that determined by the use of the series $B(x)$, we need only note the following facts:

1. Since $A_i(x)$ commences with the term $a_{11}{}^i x$, the series $U_i(x)$ must

necessarily contain no power of x less than the $(i+1)$st, as is readily verified by the use of the recursion formula for the U's.

2. The equation (4) is an identity for n, a positive integer, say $i-1$. For then,

$$\alpha_i = \frac{1}{a_{11}} \frac{(a_{11}{}^0 - 1)\alpha_{i-1}}{(a_{11}{}^i - 1)},$$

will vanish, and all succeeding α's, while the non-vanishing terms add up to exactly $A_i(x)$, as is readily verified.

3. As a consequence of the two facts just mentioned, the coefficient of x^p, in $A_n(x)$, p a positive integer, and n arbitrary, will depend upon only the first p, U's. It will therefore be a polynomial in $a_{11}{}^n$ and contain n in no other way, the coefficients being rational in a_{11} and integral, rational in the first p coefficients of $A_1(x)$.

4. The coefficients of x^p in $A_n(x)$, as obtained by the use of $B(x)$, will also satisfy the conditions just stated, and since these two polynomials in $a_{11}{}^n$ must coincide for an infinite number of values of $a_{11}{}^n$, viz. for $n = p+1$, $p+2, \cdots$, the polynomials are identical. Thus the two determinations of $A_n(x)$ coincide.

Another* and somewhat analogous representation for $A_n(x)$ may be obtained from (4), in which the variable n occurs in a more convenient form, but in which the coefficient of x^p, $p > 1$, is in general an infinite series in a_{11}. This is defined in terms of V's, as follows:

$$\begin{aligned} V_0(x) &\equiv A_0(x) \equiv x, \\ V_1(x) &\equiv A_1(x) - a_{11}A_0(x), \\ V_2(x) &\equiv [A_2(x) - a_{11}A_1(x)] - a_{11}[A_1(x) - a_{11}A_0(x)] \\ &\equiv A_2(x) - 2a_{11}A_1(x) + a_{11}{}^2A_0(x), \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

$$\begin{aligned} (5)\quad V_i(x) &\equiv V_{i-1}[A_1(x)] - a_{11}V_{i-1}(x) \\ &\equiv A_i(x) - \binom{i}{1}a_{11}A_{i-1}(x) + \binom{i}{2}a_{11}{}^2A_{i-2}(x) - \cdots \pm a_{11}{}^iA_0(x), \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

$$\begin{aligned} A_n(x) \equiv a_{11}{}^n[V_0(x) + \binom{n}{1}\frac{1}{a_{11}}V_1(x) + \binom{n}{2}\frac{1}{a_{11}{}^2}V_2(x) + \cdots \\ + \binom{n}{i}\frac{1}{a_{11}{}^i}V_i(x) + \cdots]. \end{aligned}$$

In Type IIa, $A_1(x)$ is the identity. This case requires no further men-

* Equations (4) and (5) are given here for the first time.

tion. Type IIb is a well-known form discussed by many of the authors already mentioned. By the use, for example, of the method of undetermined coefficients, the $A_1(x)$ may here be transformed through a suitable $F(x)$ into a series of the form x times a power series in x^k, the leading term being exactly x. This applies also to Type IIc. There is a wide degree of freedom in the choice of the series. In fact only the coefficients of the terms in x, x^{k+1}, and x^{2k+1} are preassigned. For Type IIb, we may in particular take the series to be in the form

$$y = \tilde{A}_1(x) = \frac{x}{\sqrt[k]{1 - k\tilde{a}_{1,\,k+1}x^k}}.$$

If in this expression we replace y by $y^{1/k}$, and x by $x^{1/k}$, we obtain

$$y = \frac{x}{1 - k\tilde{a}_{1,\,k+1}x}.$$

Replacing next, y by $1/y$, and x by $1/x$, we have

$$y = x - k\tilde{a}_{1,\,k+1},$$

so that in this way we have secured a transform of $y = A_1(x)$, in the form

$$y = c + x.$$

Putting $-k\tilde{a}_{1,\,k+1}y$ for y and $-k\tilde{a}_{1,\,k+1}x$ for x, we arrive at exactly

$$y = 1 + x,$$

as desired. We need merely retrace our steps to secure the transforming series $B(x)$.

In the case of Type IIc, the best that we can do for the first two steps is to transform $y = A_1(x)$ into

$$y = \frac{x}{1 - k\tilde{a}_{1,\,k+1}x - k\left[\tilde{a}_{1,\,2k+1} - \frac{1+k}{2}\tilde{a}_1{}^2{}_{,\,k+1}\right]x^2}.$$

This is an analytic function in the neighborhood of the origin, that resembles somewhat the analytic function

$$y = \frac{x}{1 - cx},$$

where c is a suitably chosen constant. Any transforming function, however, which carries the given function into one of the latter type, must have an essential singularity at the origin, it being understood that for Type IIc, $2\tilde{a}_{1,\,2k+1} \neq (k+1)\tilde{a}_1{}^2{}_{,\,k+1}$. For a fairly complete discussion of the existence of this transforming function, *cf.* Leau, *loc. cit.* It will be unnecessary for

us to go into this particular question, since a direct representation in series may be found without the use of a transforming function. We need only regard Case II, as a limiting form of Case I, and pass to the limit in either (4) or (5). We obtain in both cases the same series.* We shall have:

$$
\begin{aligned}
U_0(x) &\equiv A_0(x) \equiv x, \\
U_1(x) &\equiv A_1(x) - A_0(x), \\
U_2(x) &\equiv A_2(x) - 2A_1(x) + A_0(x), \\
&\cdots\cdots\cdots \\
U_i(x) &\equiv U_{i-1}[A_1(x)] - U_{i-1}(x), \\
&\equiv A_i(x) - \binom{i}{1} A_{i-1}(x) + \binom{i}{2} A_{i-2}(x) - \cdots \pm A_0(x), \\
&\cdots\cdots\cdots\cdots\cdots
\end{aligned}
$$

while for α_i we have

$$\alpha_i = \frac{n - i + 1}{i} \alpha_{i-1} = \binom{n}{2}, \qquad \alpha_0 = 1.$$

Hence, when $a_{11} = 1$, *i. e.*, for Case II,

$$\text{(6)} \quad A_n(x) \equiv A_0 + \frac{n}{1}(A_1 - A_0) + \frac{n(n-1)}{1 \cdot 2}(A_2 - 2A_1 + A_0) + \cdots.$$

For Case III, if $A_1(x)$ be given, there exist series $C(x)$, for which

$$C[A_1(x)] \equiv A_1[C(x)],$$

but these C's, instead of containing but a single parameter, contain more than one, and in some cases, contain an infinite number of parameters. For example $A_1(x) \equiv -x$ is commutative with any odd function of x. But if $C_1(x)$ be commutative with $A_1(x)$ then $C_m(x)$ is commutative with $A_n(x)$, where m and n are any positive integers, although the converse is not necessarily true. For c_{11} an integral root of unity, we may choose m and n so that $a_{m1} = 1$, and $c_{m1} = 1$. This carries us back to Case II. In particular, by taking $m = 1$, we may conclude that if $A_1(x)$ be such that a $C_1(x)$ exists for which c_{11} is equal to one, then $C_1(x)$ can contain, while c_{11} is kept equal to unity, at most one free parameter, unless $A_1(x)$ is of Type IIIa. Those of Type IIIa all contain an infinite number of parameters as we shall presently show. The same remarks apply when c_{11} is any integral power of a_{11}, say the l'th, since $C_1[A_{h-l}(x)]$ will also be commutative with $A_1(x)$, where h is the lowest positive integer for which $a_{11}{}^h = 1$, but $C_1[A_{h-l}(x)]$ will have unity for the coefficient of the first term. Apart

* This series has been obtained directly by Schroeder, Math. Ann., 2 (1870), p. 317.

from c_{11}, the only arbitrary parameter is a certain $c_{1,\,k+1}$, as is readily verified by the method of undetermined coefficients. Now x itself, is certainly commutative with $A_1(x)$, so that there exists at least one series $C_1(x)$ for which $c_{11} = 1$, $c_{1i} = 0$, $i = 2, 3, \cdots$. Since all series $C_1(x)$ with the same c_{11} coincide as far as the term containing the parameter $c_{1,\,k+1}$, it is clear that $A_h(x)$ which is commutative with $A_1(x)$, and for which $a_{h1} = 1$, is of the form

$$A_h(x) \equiv x + 0 + \cdots + 0 + a_{h,\,k+1}x^{k+1} + a_{h,\,k+2}x^{k+2} + \cdots.$$

The method of undetermined coefficients which gives the k to be used, shows us that this k is h itself, or an integral power of h. Hence had we started with a series

$$A_1(x) = x + 0 + \cdots + 0 + a_{1,\,k+1}x^{k+1} + a_{1,\,k+2}x^{k+2} + \cdots$$

where $a_{1,\,k+1} \neq 0$, then the only expressions $A_{(1/h)}(x)$, which can exist, will be for h, a factor of k. For types IIIb and IIIc, integer iterates, of course, always exist, and they will either be of the same type as the given series or of the analogous type under Case II. In some instances a few fractional iterates of the type of the given series may exist. However, no series $A_n(x)$ of Types IIIa, or IIIb, can exist, containing an arbitrary parameter, n.

Type IIIa demands further attention.* For the sake of definiteness, we shall first consider the case in which $h = 2$. Here $A_1(x)$ is such that if $y = A_1(x)$ then $x = A_1(y)$. Viewed geometrically, this gives us a curve in the x, y plane, whenever the series $A_1(x)$ is convergent. This curve is unaltered by interchanging x and y, that is, it is symmetrical about the line $y = x$. If we transform the coordinates by writing

$$u_0 = \tfrac{1}{2}(x - y),$$
$$u_1 = \tfrac{1}{2}(x + y),$$

the curve will be symmetrical about the line $u_0 = 0$, i. e., the u_1-axis, and in the neighborhood of the origin it will have contact with the u_0-axis, since the slope of the curve $y = A_1(x)$ is unity, at the origin. Thus, if in the neighborhood of the origin, u_1 be represented as a power series in u_0, for points on the curve, the series will represent an even function of u_0. Furthermore, this is the *only* restriction imposed upon u_1, as a function of u_0. We may select for u_1, an arbitrary even function of u_0, and determine x and y by the equations

$$x = \quad u_0 + u_1,$$
$$y = -u_0 + u_1.$$

* This leads to $A_h(x) = x$, which is known as "Babbage's Problem." Cf. the following: J. D. Gergonne, Ann. Math., 12 (1821–2), p. 73. O. Rausenberger, Math. Ann., 18 (1881), p. 379. E. Iaggi, Nouv. Ann. (3), 19 (1900), p. 483. M. J. van Ufen, Amst. Ak. Versl., 18 (1909), p. 860, and 19 (1910), p. 27–31. The discussion here is new.

If from these equations we solve for y as a function of x by eliminating the u_0 and u_1, the solution $y = A_1(x)$, will be such that also $x = A_1(y)$. We may forego the geometric picture, and select u_1 as an arbitrary series in even powers of u_0 without constant term, and not necessarily convergent. The series $A_1(x)$ obtained by solving as before, for y, will be determinate, and if $y = A_1(x)$, $x = A_1(y)$.

The discussion for a general integer value of h, follows along similar lines, although no simple geometric illustration of an obvious sort suggests itself. We shall first suppose a series $A_1(x)$ given for which $A_h(x) \equiv x$. From this series we shall determine h series u_i, $i = 0, 1, \cdots, h-1$, which we shall find to be restricted in a very simple manner only. This done, we shall start with the h series u_i and so determine the form of the most general series $A_1(x)$ for which $A_h \equiv x$. Let us write

$$u_0 = \frac{1}{h}(A_0 + \omega A_1 + \omega^2 A_2 + \cdots + \omega^{h-1} A_{h-1}),$$

$$u_1 = \frac{1}{h}(A_0 + \omega^2 A_1 + \omega^4 A_2 + \cdots + \omega^{2h-2} A_{h-1}),$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$u_{h-1} = \frac{1}{h}(A_0 + A_1 + A_2 + \cdots + A_{h-1}),$$

where ω is a primitive hth root of unity. We wish now to discover what restrictions must obtain among the u's in order that $A_h \equiv A_0$. If we replace A_i by A_{i+1}, u_0 is carried into $\omega' u_0$, where ω' is the reciprocal of ω, while u_{h-1} is unaltered. Hence if we express u_{h-1} as a power series in u_0, by eliminating the independent variable x, occurring in each A_i, $i = 0, 1, \cdots, h-1$, we shall find that only the terms in powers of u_0^h appear, for if any other term were to appear, it would be changed into the product of its original value by a power of ω, different from unity, which is impossible, since u_{h-1} is to remain unaltered. In a similar manner the power series for u_i, $i = 1, 2, \cdots, h-1$, in terms of u_0, contain only terms in which u_0^{i+1} times a power of u_0^h appears. These are the only restrictions upon the series u_i in terms of u_0. We may suppose, if we desire, that the u's are given, and that we are to find $A_1(x)$. We may take any $h-1$ power series, $P_i(t)$, $i = 1, 2, \cdots, h-1$, convergent or divergent and write $u_i = u_0^{i+1} P_i(u_0^h)$. If we solve for the A's formally in terms of the u's, we obtain in particular

$$A_0(x) = x = (u_0 + u_1 + \cdots + u_{h-1})$$

i. e., $x = E_{-1}(u_0)$ where $E_{-1}(u_0)$ is, by definition, the series in u_0 obtained by replacing each u_i, $i = 0, 1, 2, \cdots, h-1$ by its series representation in

terms of u_0, in the right hand member of the above equation. By definition of E_{-1}, it follows that

$$u_0 = E(x).$$

Now $E_{-1}(u_0)$ is of the form

$$E_{-1}(u_0) \equiv u_0 + u_0^2 P_1(u_0^h) + u_0^3 P_2(u_0^h) + \cdots + u_0^h P_{h-1}(u_0^h).$$

Thus $E_{-1}(\omega u_0)$ is simply $(\omega u_0 + \omega^2 u_1 + \omega^3 u_2 + \cdots + u_{h-1})$, since when we replace u_0 in $u_i = u_0^{i+1} P_i(u_0^h)$, by wu_0, we obtain merely $\omega^{i+1} u_i$. But this expression for $E_{-1}(\omega u_0)$ is exactly what we obtain formally, when we solve for A_1 in terms of the u's. Thus we have $A_1 \equiv E_{-1}(\omega u_0)$ or

$$A_1(x) \equiv E_{-1}[\omega E(x)],$$

from which we verify that indeed $A_h(x) \equiv x$, since obviously for i any positive integer we have

$$A_i(x) \equiv E_{-1}[\omega^i E(x)].$$

Thus we verify that it is not only necessary but sufficient for $u_i(u_0)$ to be of the form $u_0^{i+1} P_i(u_0^h)$, in order that the $A_1(x)$ associated with the u's as above shall be such that $A_h(x) \equiv A_0(x)$. The $E_{-1}(u_0)$ that we have obtained above is a power series in u_0 characterized by the property that the absolute term is missing, the coefficient of u_0 is unity, and the coefficient of u_0^{lk+1}, $l = 1, 2, \cdots$, vanishes. For any given series $A_1(x)$ the series E_{-1} satisfying the conditions just mentioned is uniquely determined by the equation

$$A_1(x) \equiv E_{-1}[\omega E(x)]$$

analogous to the equations obtained in Case I. Unlike the situation in Case I, we have here imposed additional conditions on E_{-1}. If we replace u_0 by a power series $u_0 = S_{-1}(v)$, where S_{-1} is restricted to the extent of having the coefficient of every term which does not contain v to a power of the form $lk + 1$, $l = 0, 1, 2, \cdots$, vanish, while the remaining coefficients are arbitrary, then we obtain

$$A_i(x) \equiv E_{-1}\{S_{-1}[\omega^i S\{E(x)\}]\},$$

where A_i, E, and ω are fixed and S_{-1} has an infinite number of arbitrary coefficients as already remarked. We may think of $E_{-1}[S_{-1}(v)]$ as constituting a new series $F_{-1}(v)$, for which the coefficients of v^{lk+1}, $l = 1, 2, \cdots$ are arbitrary, the others being then determined, if A_1 be given, so that

$$A_1(x) = F_{-1}[\omega F(x)].$$

Case IV presents little of interest. Type IVa is trivial, and will be passed over without further comment. Type IVb has a general iterate as a function of n, but it is not expressible as a power series in integral powers

of x; terms with irrational exponents, for example, must be introduced. We may however obtain a transforming function, which reduces this case to Case I. We first choose α, as one of the $(k-1)$st roots of a_{1k}, and replace y by y/α, and x by x/α. We obtain in this way a new series of Type IVb, but one in which a_{1k} is now unity. By the method of undetermined coefficients, we may find a series $F(x)$ such that

$$F[A_1(x)] \equiv [F(x)]^k,$$

so that $A_1(x)$ is now transformed into x^k. The equation $y = x^k$ is transformed into $y = kx$, by replacing y by e^y and x by e^x, while $y = kx$ is of the form Type Ib.

The Transforming Series as a Limit Series.

8. We shall now suppose that an $A_1(x)$ and a $G_1(x)$ are given and known, while the existence of an $F_1(x)$ is known but not the form of the series, where

$$A_1(x) \equiv F\{G_1[F_{-1}(x)]\}.$$

We shall also suppose that the general iterate $G_n(x)$ is known, and we wish to determine F so that

$$A_n(x) \equiv F\{G_n[F_{-1}(x)]\}.$$

We shall now describe a method which in some important cases serves to define an $F(x)$. The method is not always applicable but is of considerable interest in the cases in which it can be used.*

For any given positive integer p, let us consider the four series $G_{-p}[A_p(x)]$, $G_p[A_{-p}(x)]$, $A_{-p}[G_p(x)]$, $A_p[G_{-p}(x)]$, of which the latter two are the inverses respectively of the former two. A given term in any one of these series will be a function of p, which may or may not remain finite as p takes on successively larger and larger positive integer values. If it should happen that in one of these series the coefficient of each term approaches a definite finite limit for $p = +\alpha$, then a limit series, $L(x)$ will be defined. The question is one of convergence, not of the series as a series in x, but of the coefficients of the series as functions of p. If $L(x)$ exists, its inverse also exists, so that we may without loss of generality suppose L to be the limit series of one of the first two mentioned above. From the definition of $L(x)$ it then follows that

$$G[L(x)] \equiv L[A(x)],$$

so that $L(x)$ is a series of the form $F_{-1}(x)$ which was to be found. In determining $L(x)$ we have not made use of the actual form of $G_n(x)$, for n other than an integer, so that when different methods of interpolation are used to give different forms for $G_n(x)$, correspondingly different forms will

* Cf. G. Koenigs, Ann. Ec. Norm. (3), 1 (1884), Suppl., p. 19, and (3) 2 (1885), p. 385.

be obtained for $A_n(x)$, n other than an integer. If $G_{-p}[A_p(x)]$ and $G_p[A_{-p}(x)]$ both have limit series L, and these are distinct, then for every integer value of n, we shall have

$$G_n[L(x)] \equiv L[A_n(x)]$$

true, for both choices of $L(x)$, but for fractional, irrational or complex values of n, it may well happen that the same determination of $G_n(x)$ gives for the two L's different determinations of $A_n(x)$, which corresponds to the arbitrary feature of $G_n(x)$. If $C(x)$ be any series commutative with $G(x)$, then $C[L(x)]$ may be used equally well, in place of $L(x)$. In particular $C(x)$ may be an integral or any fractional iterate of $G(x)$. But this corresponds merely to having considered $G_{-p+q}[A_p(x)]$ or $G_{p-q}[A_{-p}(x)]$, respectively, in place of the given series, where q is a fixed number. Similarly we might have taken $L[C(x)]$ in place of $L(x)$, where $C(x)$ is this time commutative with $A(x)$. These two cases are not, of course, independent, since, as we have already remarked, $G_n[L(x)] \equiv L[A_n(x)]$. The one-parameter family of series $G_q[L(x)]$, q arbitrary, which exists if $L(x)$ does, constitutes in many cases the totality of series $F_{-1}(x)$.

Investigation of the Convergence of the Series Employed.

9. We now come to the question of convergence. In Type Ia, whenever $A_1(x)$ is a convergent series, the convergence of $E(x)$ follows, and conversely, as may be proved in several ways.* Type Ib, although not usually treated, presents no difficulties, since $A_1(x)$ and $A_{-1}(x)$ yield the same $E(x)$, and since when $|a_{11}| > 1$, then $|a_{-1,1}| < 1$, $a_{-1,1}$ being merely $1/a_{11}$. Thus, if $A_1(x)$ is convergent and $|a_{11}| > 1$, then $E(x)$ is convergent, because of the convergence of $A_{-1}(x)$. Type Ic is of a different sort. For $A_1(x)$ convergent and of Type Ic, $E(x)$ may be either convergent or divergent. If $A_1(x)$ is convergent certain conditions must be satisfied in order that $E(x)$ be convergent. These are not formal conditions upon the first p coefficients, $p = 1, 2, \cdots$, since the problem is one having significance only for convergent series $A_1(x)$, and convergence is not dependent upon the values of any finite number of coefficients. If $E(x)$ be convergent, $A(x)$ must be. Necessary and sufficient conditions for the convergence of $E(x)$, $A_1(x)$ being given as convergent, may be given in different forms. We shall state a condition which theoretically is capable of explicit application in any case. The cases of Type Ic, for which $E(x)$ is convergent, we shall call of Type Ic', those for which $E(x)$ is divergent, being of Type Ic'', provided that $A_1(x)$ is convergent, otherwise of Type Ic'''. We shall now state

* For example, Leau treats the equation in this instance both by the method of dominant functions, and by that of successive substitutions, loc. cit.

A necessary and sufficient condition for Type Ic′. It must be possible to find a circle about the origin in the complex plane such that if any point P other than the origin be taken within this circle, the totality of the positive and negative iterates of P, constitute a set of points lying everywhere densely upon a simple closed curve of finite length, and with continuously turning tangent, encircling the origin.

This includes, in particular, the conditions of interior and exterior "stability," about the origin. The condition of interior stability, frequently called merely "stability," is the following: For any sufficiently small circle C_1 with center at the origin, there must be a smaller concentric circle C_2, such that the positive and negative iterates of every point within C_2 is within C_1.* The condition for exterior stability insures that every iterate of a point in the neighborhood of the origin but exterior to a small circle enclosing the origin remains exterior to the same or a smaller fixed circle about the origin.

To prove that the condition stated above is sufficient, we may notice, that if it be satisfied for a given P, we shall have a curve C completely determined by the iterates of P, together with the conditions that C is continuous, and its tangent turns continuously. This curve C may be transformed into a circle in the plane of a complex variable t by a transformation of the form $t = E_{-1}(x)$, where $E_{-1}(x)$ is an analytic function, as follows from the theory of conformal mapping. The function E_{-1} will be completely determined when we require the origin in the x-plane to go into the origin in the t-plane, the circle in the t-plane being the unit circle and the positive real axis in the x-plane going into some sort of curve tangent to the positive real axis in the t-plane, at the origin. Any analytic transformation of the x-plane in the neighborhood of the origin, which carries the curve C into itself and the origin into itself, will correspond analytically to a transformation of the t-plane in the neighborhood of the origin, leaving the unit circle and the origin unaltered. Now $A_1(x)$ is such a transformation of the x-plane, and it corresponds therefore to a rotation in the t-plane, since a rotation is the only conformal transformation of the unit circle into itself which leaves the center invariant.† But formally it corresponds to a rotation $a_{11} \cdot t$, $|a_{11}| = 1$, where a_{11} is the initial coefficient of the series $A_1(x)$. Thus the formal series $E(x)$ is actually the inverse of the series corresponding to the analytic function $E_{-1}(x)$ just determined. Hence the above condition is sufficient; that it is necessary is seen immediately from the fact that if E is analytic then the successive iterates of a point P in the x-plane correspond to the successive images of a point in the t-plane, as the t-plane is rotated through an angle incommensurable with 2π.

* Cf. Levi Civita, Ann. di. Mat. (3), 5 (1901), p. 240, and Cigala, Ann. di. Mat. (3), 11, (1905), p. 67.

† Cf. Osgood, Funktionentheorie, vol. I, 1st ed., p. 595–.

In the case of Type II*a*, we may transform x into itself by the use of divergent as well as of convergent series. No simple or direct proof of the convergence of the series transforming Type II*b* into

$$y = \frac{x}{1 - 2\tilde{a}_{12}x},$$

seems to have been made. The convergence follows, however, when $A_1(x)$ is itself, convergent, from the detailed investigation of a function-theoretic character made by Leau.* Since the series used in Type II*c* are not formally equivalent to

$$y = \frac{x}{1 - 2\tilde{a}_{12}x},$$

no convergent power series can be found which transforms the former into the latter. The existence of functions in general analytic, but possessing essential singularities at the origin is proved by Leau for the cases in which the given series is convergent.

The existence of divergent series of the Type III*a* has already been shown. Whenever the series is convergent, the h series, u_i, $i = 0, 1, \cdots, h - 1$, are convergent, since they are linear combinations of the first h iterates of $A_1(x)$, so that the corresponding $E(x)$ must be necessarily convergent, while for $A_1(x)$, divergent, $E(x)$ is divergent. In the cases of Types III*a* and III*b*, for which, in general, no iterate containing an arbitrary parameter n, can exist of Case III, it is not surprising that no transformation is known which reduces either of these types to the form cx, or $x + a_0$, and certainly no conformal transformation of this sort can exist, so that there is no question of convergence.

A discussion of the convergence of the transforming function E, which reduces Type IV*b*, to the form $y = x^k$ has been made by Grèvy, *loc. cit.*

A Geometric Interpretation of Iteration.

10. We have already mentioned an infinite matrix whose first principal minor of order m, we have denoted by $\boldsymbol{A}_1^{(m)}$, which may be considered as serving to define the series $A_1(x)$. We have mentioned the problem of the reduction of this infinite matrix to a normal form by successively reducing to normal form the matrices $\boldsymbol{A}_1^{(m)}$, $m = 1, 2, \cdots$. In place of the collineation symbolized by $y = \boldsymbol{A}_1^{(m)}x$, and its reduction to a normal form, we may consider the entirely equivalent problem of the simultaneous linear transformation of the pair of bilinear forms, or of the λ-matrix, corresponding to the original matrix. In terms of the original series, $A_1(x)$, this means that we may consider the analytic transformation of the pair of

* Leau, loc. cit.

expressions $y = x$, and $y = A_1(x)$, or of the linear function of λ, expressed by $y = A_1(x) - \lambda x$. Since any analytic curve in the real plane may be transformed analytically in the neighborhood of a non-singular point of the curve into $y = x$, the invariants under analytic transformations of a pair of curves at a point of intersection, are the invariants of the pair $y = x$, and $y = A_1(x)$, where $A_1(x)$ is the transform of the second curve. This excludes, of course, certain exceptional cases, as when a derivative of $A_1(x)$ becomes infinite, which, however, may be readily treated as limiting cases. The group of transformations, $y' = \varphi(y)$, $x' = \varphi(x)$, has significance chiefly in the "function-plane," and by change of variables, in the "inversion-plane." The study of the conformal invariants of a pair of real curves at a real point of intersection, is essentially the study of Cases II and III, discussed above, while the general conformal group without regard to reality involves also Cases I and IV.* Integral iteration corresponds to the multiplication of the angle between $y = A_1(x)$ and $y = x$, while division of the angle is secured by fractional iteration.

PART TWO. THE ITERATION OF A REAL FUNCTION.

The Domain of Definition of A_n.

11. We shall consider the iteration of a real function defined for the points within an interval a, b, $a < b$, where, in particular, a may be negatively infinite, or b positively infinite, or both. For many problems in analysis, it is immaterial whether we consider a real analytic function as defined also for imaginary points, or not, so far as the phenomena for real points alone are concerned. Iteration is not, however, one of these problems, since it frequently happens that the iterate of an imaginary point is itself a real point.

Let $A_1(x)$ be a finite, real, one-valued function of x, defined over a range r_1 in the interval a, b. We shall usually suppose that the range r_1' of values taken on by $x' = A_1(x)$ lies within the interval a, b, since otherwise, we might have extended our original interval a, b, so as to include r_1', without altering r_1 or any other data.

The function $A_2(x) \equiv A_1[A_1(x)]$, will be defined only for the numbers of the subset r_2 common to r_1 and r_1'. If we denote by r_2', the values assumed by $A_1(x)$ as x varies over r_2, then $A_3(x)$ is defined only for the subset r_3, common to r_2 and r_2', and so forth. It may happen that after iterating a finite number of times, we may eventually find an n, so large that r_n is a null-set, and no expression $A_n(x)$ can be defined. This case, although common, presents but little of interest. If r_1 be an open set, it may happen

* Cf. Kasner, "Conformal Geometry," loc. cit., and Pfeiffer, notice in Bull. Am. Math. Soc., 21 (1915).

that for every positive integer n, r_n contains some number, while no number is common to all r_n's, owing to the fact that the limits approached lie outside of r_1. A third case arises when the limit set, r_∞, consists of a finite number of points, $e_1, e_2, \cdots, e_k$, so that these are the only points for which $A_n(x)$ can be defined for every positive integer n. Now $A_1(e_i)$ must be equal to one of the e's, say e_j, for if $A_n(e_i)$ is defined for all positive integral values of n, so must also $A_{n-1}[A_1(e_i)]$, since these are identical by definition, or in other words $A_1(e_i)$, is an e by definition. Since $A_1(x)$ is one-valued, each e_i must go into a unique e_j either itself or another. Of the k, e's, $e_1, e_2, \cdots, e_k$, it may happen that several distinct e's are carried into a single one by the operation $x' = A_1(x)$. There must be a non-vanishing subset of these k, e's into each one of which subset, an e of the original set is carried by $x' = A_1(x)$. By repeated selections of such subsets, we obtain finally a non-vanishing set of k', e's, which we may call $e_1, e_2, \cdots, e_{k'}$, such that each e of this set is carried into an e of this set and is obtained from an e of this set by any transformation $x' = A_n(x)$, n, a positive integer. Since k' is finite, and each e_i is carried into but a single e_j, each e_j must be obtained from but a single e_i of the set of k', e's. We have then a permutation group upon k' elements, and the transformation $x' = A_{k'!}(x)$ must leave each of this set unaltered. Other cases arise in which r_∞ coincides with r_1. It may possess no fixed points, a finite number, or an infinite number, and every point of r_∞ may be fixed.

We may distinguish in all cases, as we have already done in one case, a subset of r_∞, which we shall call r_∞', where not only is $A_n(x)$ defined for every point of r_∞', for n an arbitrary positive integer, but $A_{-n}(x)$ is also defined for all positive integral values of n. For the subset of r_∞ obtained by removing r_∞', we have points which are carried into other points by A_n, but which are not obtainable from other points by an A_n, where n is a sufficiently large positive integer, in each case. The subset denotable by $r_\infty - r_\infty'$, may be further subclassified. Regarded as a set of points it is completely arbitrary in character. The set r_∞', may be divided into subsets, $r_{\infty 1}', r_{\infty 2}', \cdots$, together with certain residual sets $r_{\infty}'_{,\,\mathrm{ell}}$, $r_\infty'_{\mathrm{par}}$ and $r_\infty'_{,\,\mathrm{hyp}}$. The set $r_\infty'_p$, for p a positive integer, is defined as a set which is carried pointwise into itself by $A_p(x)$, but by no $A_n(x)$, $n < p$. The set $r_\infty'_{,\,\mathrm{ell}}$ is carried into itself circularly, but no point is carried into itself exactly by any A_n, n, a positive integer. This corresponds to the case of a rotation through an irrational angle. The set $r_\infty'_{,\,\mathrm{hyp}}$, is a set carried into itself by $A_n(x)$, n a positive integer, but not pointwise, there being for any particular case two limit points, after the manner of a hyperbolic linear fractional transformation. Similarly any particular set $r_\infty'_{,\,\mathrm{par}}$ has one limit point. The set $r_{\infty 1}'$ is absolutely arbitrary in form. The set $r_{\infty 1}'$ is

restricted to the extent that it must consist of a set of pairs of points, without repetition, or as it may be regarded, as a pair of sets of points, the points in one set being in one to one reciprocal correspondence with those of the other. Similarly the set $r_{\infty n}'$ must consist of sets of n points without repetition, or n sets in one to one reciprocal correspondence with one another. The sets, $r_{\infty}'{}_{,\,\mathrm{ell}}$, may be regarded as constituting independent systems, as also in the cases of $r_{\infty}'{}_{,\,\mathrm{par}}$ and $r_{\infty}'{}_{,\,\mathrm{ell}}$.

Some General Theorems on A_n.

12. We shall state a few of the numerous theorems which are valid for real one-valued functions defined over a range r in an interval a, b, $a < b$, extending, in particular cases, to infinity. The method of proof is fairly obvious in every case, and as the proof itself is never difficult, it will be omitted. The index n, of iteration will be restricted to positive integers. We shall use the notation c, d_1 for an interval, $c \leqq x \leqq d_1$ included within the given interval.

If $A_1(x)$ is periodic, $A_n(x)$ is periodic in so far as it is defined.

Every point of intersection of $y = A_1(x)$ with $y = x$, is also a point of intersection of $y = A_n(x)$ with $y = x$.

If $c_1 \leqq A_1(x) \leqq d_1$, whenever $c_2 \leqq x \leqq d_2$, for $c_2 \leqq c_1 < d_1 \leqq d_2$, then the same is true for $A_n(x)$.

If $0 \leqq s_1 < s_2$, and $A_1(e) = e$, and if also for x in the interval c, d, where $c \leqq e \leqq d$, $A_1(x)$ lies between $e + s_1(x - e)$ and $e + s_2(x - e)$, then for x in the same interval, $A_n(x)$ will lie between $e + s_1{}^n(x - e)$ and $e + s_2{}^n(x - e)$

If e be a value for which $A_1(e) = e$, then every point of intersection of $y = A_1(x)$ and $y = e$ is also a point of intersection of $y = A_n(x)$ and $y = e$.

Hence we obtain also the following theorem,

If $c \leqq A_1(x) \leqq d$, for $a \leqq x \leqq b$, and if $y \equiv x$ for $c \leqq x \leqq d$, then $A_n(x) \equiv A_1(x)$.

If for every $x_1 < x_2$, $A_1(x) < A_1(x)_2$, then also $A_n(x_1) < A_n(x_2)$, in so far as $A_n(x)$ is defined.

If the absolute minimum of $A_1(x)$ is unique and occurs at c, and $A_1(c) > c$ and also the absolute maximum of $A_1(x)$ is unique and occurs at d, and $A_1(d) < d$, and if $A_1(x)$ in so far as it is defined is an increasing function for $c \leqq x \leqq d$, then whenever $A_n(c)$ is defined it will be the unique absolute minimum value of $A_n(x)$, and $A_n(c) > A_{n-i}(c)$, $i = 1, 2, \cdots, n$, in so far as these are defined. A corresponding situation holds for $A_n(d)$, and $A_n(x)$ will be, in so far as it is defined, an increasing function of x for $c \leqq x \leqq d$.

If for every $x_1 < x_2$, $A_1(x_1) > A_1(x_2)$, then also $A_{2n-1}(x_1) > A_{2n-1}(x_2)$, while $A_{2n}(x_1) < A_{2n}(x_2)$ in so far as $A_{2n-1}(x)$ and $A_{2n}(x)$ are defined.

If $A_1(a)$ is either undefined or is equal to c_1 and for $a \leqq x_1 < x_2 \leqq c_2$, $c_1 \leqq A_1(x_1) < A_1(x_2) \leqq b$, and if for $c_2 \leqq x_1 < x_2 \leqq b$, $a \leqq A_1(x_1) < A_1(x_2) \leqq c_1$, while $A_1(b)$ is either undefined or equal to or less than c_1, then there exists a pair of number $c_1^{(n)}$ and $c_2^{(n)}$, for which an analogous situation holds for $A_n(x)$.

If in the preceding theorem we put "$\geqq$" in place of "$\leqq$", and $2n - 1$ for n, the theorem continues to hold, but if we put "$\geqq$" in the hypothesis and $2n$ in the conclusion, we must read "$\leqq$" in the conclusion.

Other theorems might be mentioned of a nature analogous to those just cited. We shall now however give a few of a different type, in which the continuity of $A_1(x)$ is supposed, as well as the existence and in some cases the continuity of the first derivative of $A_1(x)$ with respect to x. The proofs are obvious and will be omitted. They involve chiefly the use of the identity,

$$A_n'(x) = A_1'[A_{n-1}(x)] \cdot A_1'[A_{n-2}(x)] \cdots A_1'[A_1(x)] \cdot A_1'(x),$$

where $A_n'(x)$ denotes $\dfrac{\partial A_n(x)}{\partial x}$.

If the total number of maxima of $A_1(x)$ be finite and equal to m, while $A_1'(x)$ does not remain zero over any interval, the least possible number of maxima of $A_n(x)$ will be $m - 1$, and the greatest possible number will be $2^{n-1}m^n$. Examples of the two cases are most readily grasped by means of figures. Fig. 1 represents a broken line joining in order the points whose

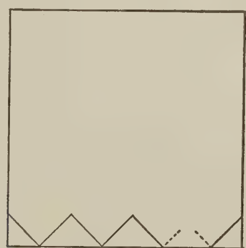


FIG. 1.

coordinates are, $[0, 1/(m - 1)]$, $[1/(m - 1), 0]$, $[2/(m - 1), 1/(m - 1)]$, $[3/(m - 1), 0]$, $\cdots$, $[2 - 1/(m - 1), 0]$, $[2, 1/(m - 1)]$. Odd iterates of

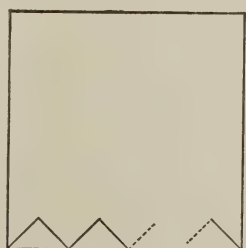


FIG. 2.

this curve coincide with the curve itself while even iterates are of the form illustrated in Fig. 2, where we have a broken line joining in order the points whose coordinates are $[0, 0]$, $[1/(m-1), 1/(m-1)]$, $[2/(m-1), 0]$, $[3/(m-1), 1/(m-1)]$, $\cdots$, $[2 - 1/(m-1), 1/(m-1)]$, $[2, 0]$. An example of the second case is given in Fig. 3, by the broken line joining in

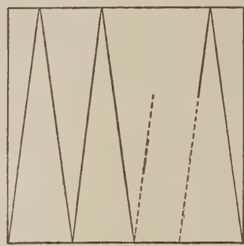


FIG. 3.

order, the points, $[0, 0]$, $[1/m, 2]$, $[2/m, 0]$, $[3/m, 2]$, $[4/m, 0]$, $\cdots$, $[2 - 1/m, 2]$, $[2, 0]$. The above theorem holds also if " minimum " be read for " maximum."

If for every x, $| A_1'(x) | < s$, where s is a constant, than $| A_n'(x) | < s^n$, throughout the interval. Hence if $y = A_1(x)$, be a curve of limited variation, so also is $y = A_n(x)$. The length of arc of $y = A_n(x)$ between $x = c$ and $x = d$, $c < d$, will be less than $(d - c)s^n$.

If $0 < s_1 < A_1'(x) < s_2$, then $0 < s_1^n < A_n'(s) < s_2^n$.

If $| A_1'(x) | \leqq 1$, throughout the interval, then the number of maxima of $A_n(x)$ cannot differ by more than one from the number of maxima of $A_1(x)$, provided that $A_1'(x)$ does not remain equal to zero over an interval. The same theorem holds true if " minima," be read for " maxima."

If throughout the interval, $A_1(x) \leqq a + (x-a)^k/(b-a)^{k-1}$, then $A_n(x) \leqq a + (x-a)^{k^n}/(b-a)^{k^n-1}$. The same theorem holds if "$\leqq$," be replaced by "$\geqq$."

One-Valued Iteration, A_n, for n Varying Continuously.

13. Let us examine those cases in which $A_1(x)$ is one-valued, and $A_n(x)$ exists as n varies continuously over real values. We shall have given a range r over which $A_1(x)$ is defined. In r, we shall distinguish two classes of points, which we shall call *singular points*, and *ordinary points*. An ordinary point will be defined as a point of r, for which there is no next point of r whether to the right or to the left of it, the other points are the singular points. A point which is not of r, but which is a limit point of points of r, and for which on both sides, there is no next point of r, will be called a *gap point* of r. We shall suppose that except in the neighborhood of $x = a$, and $x = b$, $A_n(x)$ exists for all values of n for which $| n |$ is suffi-

ciently small. Since for n irrational, $A_n(x)$, is defined only by means of interpolation, we shall be forced to regard $A_n(x)$ for any given x in r, as varying continuously over r, when n varies continuously over real values. This does not mean that $A_n(x)$ for a given x varies continuously in the interval a, b, as n varies continuously, but simply that it varies continuously with reference to the inner structure of r. Since continuous variation from a singular point is impossible, we must have $A_n(x) = x$, for every singular point. It must be borne in mind that a point is singular or not, purely in terms of the internal structure of r, and the end point of an interval c, d of points of r need not be a singular point. Since by continuous variation in r, a gap point may be approached but not passed, we see that the limit points approached by $A_n(x)$, for x given, and n increasing by integers toward infinity include in general gap points. If we take three numbers, $x_1 < x_2 < x_3$, then for $|n|$ small, we must have $A_n(x_1) < A_n(x_2) < A_n(x_3)$, provided, at least, that $a < x_1$ and $x_3 < b$, since a continuous group of transformations of the form here considered cannot include a transformation by which sense is altered, except possibly in the neighborhoods of a and b, for $|n|$ small. We may make three possible conventions with regard to the end points a and b, which will be in accord with the conditions imposed upon interior points.

First. We may require $A_n(x)$ to be defined at a and b only when n is positive for one of these points and negative for the other.

Second. We may require $A_n(x)$ to be defined at both a and b for $|n|$ small, and for n both positive and negative.

Third. We may regard the interval a, b, as treated cyclically, so that a and b are not distinguished, and a point to the right of a is treated as to the right of any point to the left of b, only cyclic order being essential. In this case, we may, if desired, require, in particular, that r coincides with the whole interval a, b, and $A_n(x)$ is defined for every point of the interval, whenever $|n|$ be small, whether n be positive or negative.

In the first case, a and b will be essentially different from any other points of r, if they, or either one, be in r. The range of definition for $A_n(x)$, will in general depend upon n, and various small difficulties arise. It is however a common and important case. In the second case, a and b, if in r, are merely singular points, as already defined, and are not to be distinguished from the other singular points. In the third case, a and b may be regarded as a single point in no way distinguishable from any ordinary point or r, either in terms of the structure of r itself, or in terms of the behavior of $A_n(x)$, as a function of n.

We may describe the curve $y = A_1(x)$, where $A_1(x)$ satisfies the conditions of this section as being *with respect to* r, a series of monotonically

increasing curvilinear segments, and discrete points. The discrete points are all on the line $y = x$, and have for their abscissas, the singular points of r, measured on the x-axis. The curvilinear segments, if they have end points, have these on the line $y = x$. If we require that the range of definition interior to a, b, with the possible exception of the neighborhoods of a and b themselves, shall not vary with n, then we have also that curvilinear segments, where there is no end point, approach, i. e., at a gap point, the line $y = x$, and have a point of $y = x$ for limit point. When referred to the interval a, b, without the intervention of r, a more complicated description of $y = A_1(x)$, is, of course, necessary.

The Transforming Function $B(x)$.

14. In defining $A_n(x)$ for n, a fraction, we may make use of a transforming function, $B(x)$, such that if $y = A_n(x)$

$$\begin{cases} x = B(t), \\ y = B(n + t), \end{cases}$$

so that $A_n(x) \equiv B[n + B_{-1}(x)]$. If $y = A_n(x)$ is to be a one-valued function of x, for positive values of n, then $B(t)$ must be monotonic if it be continuous, that is, if $A_n(x)$ be a continuous function of n for a given x. For suppose if possible that $B(t)$ is one-valued and continuous but not monotonic. Then let $B(\alpha + t_1) = B(t_1)$, whence since $A_\alpha[B(t_1)] \equiv B(\alpha + t_1)$, we have $A_\alpha[B(t_1)] = B(t_1)$ and therefore $A_{m\alpha}[B(t_1)] = B(t_1)$, where m is any positive integer, or $B[m\alpha + t_1] = B(t_1)$, $m = 1, 2, \cdots$. $B(t)$, being supposed non-monotonic, will have either a maximum or a minimum, if not many of each. Let us suppose there is a maximum $\beta = B(\gamma)$. A line $x = \beta - \epsilon$, $\epsilon > 0$, will cut the curve, $x = B(t)$, at two points in the neighborhood of (γ, β), which we may call $t = \gamma - \delta_1$, and $t = \gamma + \delta_2$, $\delta_1 > 0$, $\delta_2 > 0$. Making use of our previous remark, we conclude that

$$B[m(\delta_2 + \delta_1) + \gamma - \delta_1] = B(\gamma - \delta_1), \qquad m = 1, 2, \cdots.$$

Thus the line $x = \beta - \epsilon$ contains a set of points of the curve $x = B(t)$ beginning with $t = \gamma - \delta_1$, and succeeding each other at a constant interval $\delta_1 + \delta_2$. As ϵ is taken smaller and smaller, we obtain new lines and each contains points of the curve $x = B(t)$ distributed in the same fashion at constant intervals from each other, the constant interval, however, decreasing with ϵ. From the continuity of B we would conclude that in the limiting case of the line, $x = \beta$ every point of this line to the right of $t = \gamma$, ought to be a point of the curve $x = B(t)$. But if $x = B(t)$ is to be one-valued, such a situation is impossible. We therefore conclude, that for $x = B(t)$, continuous and one-valued, this curve is also monotonic. The proof for

the case of a minimum but no maximum follows in the same fashion and is therefore supposed completed.

The range of the variable t, has, of course, nothing to do with the range r or the interval a, b. If $B(t)$ is monotonic and one-valued, its most significant characteristics will be its vertical and horizontal asymptotes if one or both of these exist. If $u = B(t)$ has a horizontal asymptote $u = e$, then $x = e$, is a fixed point for $y = A_n(x)$. If $u = B(t)$ has a vertical asymptote, then $y = A_n(x)$ has for every n, a fixed common horizontal asymptote, and the interval a, b must be infinite, while if the curve $u = B(t)$ be defined to the left of a vertical asymptote, then $y = A_n(x)$ has a vertical asymptote. Further theorems of a like sort are readily obtained, for instance: If $u = B(t)$ approaches minus infinity, as t approaches e, from the left, then $y = A_n(x)$, is an increasing function with a vertical asymptote $x = e - 1$, and lying to the right of the asymptotes.

A transforming function $B(t)$ must be itself monotonic and is adapted for the study of only monotonic functions $y = A_1(x)$. If in the neighborhood of a point $x = y = e$, $y = A_1(x)$ is monotonic and increasing, we may obtain for this neighborhood a transforming function $x = B(t)$. The question arises as to whether if we continue $A_1(x)$ beyond the neighborhood in which it is monotonic, there will not be a corresponding extension of $B(t)$, so that the $B(t)$ so extended shall still serve as the transforming function of $y = A_1(x)$ when this latter is extended. We shall find that no extension of $B(t)$ can be made to serve as a transforming function for $A_1(x)$ when extended beyond the region in which $A_1(x)$ is monotonic. If, however, $A_1(x)$ is analytic, we may determine $B(t)$ analytically, and each will be continued analytically beyond the region in which it is monotone. What then is the relation between these continuations? For $y = A_1(x)$, one-valued, but not monotonic, the extension beyond the monotonic interval containing a point $x = y = e$, leads us to an extension of $B(t)$ into an *auxiliary* part, such that x and y may both vary on the *principal* part $B(t)$ with the relation $x = B(t)$, $y = B(1 + t)$, but x alone varies on the auxiliary part $x = \tilde{B}(t)$, while y continues to satisfy $y = B(1 + t)$. The situation may readily be grasped upon inspecting the case in which $A(x) \equiv \sin x$. The curve $u = B(t)$ is indicated, Fig. 4, by a heavy curve, while the

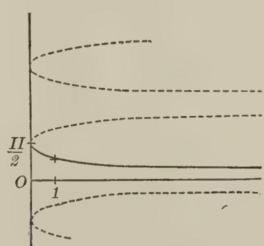


FIG. 4.

auxiliary part $u = \tilde{B}(t)$, here consists of an infinite sequence of branches a few of which are here indicated by dotted curves.

No inverse $A_{-1}(x)$, in a strict sense, exists when $A_1(x)$ is non-monotonic. For example, the equation $y = \text{arc sin}\,[\sin x]$, is satisfied not only by $y = x$, but also by $y = x + 2m\pi$, and $y = (2m + 1)\pi - x$, where m is an arbitrary integer. For these reasons the study of the complete behavior in the complex plane of a transforming function $B(t)$, for a given $A(x)$, does not present the interest that might be at first expected.

Multiple-Valued Functions. Weights.

15. We shall now consider the case of multiple-valued real functions. Let $y = A_1(x)$, $y = F_1(x)$, $y = G_1(x)$, $\cdots$, be a set of single-valued functions, which taken in their totality constitute a multiple-valued function $y = \varphi_1(x)$. The expression $y = \varphi_2(x)$ may be regarded as an abridged way of writing the set of one-valued functions

$$\begin{array}{llll} y = A_2(x), & y = A_1[F_1(x)], & y = A_1[G_1(x)], & \cdots, \\ y = F_1[A_1(x)], & y = F_2(x), & y = F_1[G_1(x)], & \cdots, \\ y = G_1[A_1(x)], & y = G_1[F_1(x)], & y = G_2(x), & \cdots. \end{array}$$

In most cases of multiple-valued functions φ, the expression, $y = \varphi_n(x)$, can have no meaning for n other than a positive integer. It sometimes happens that two or more of the original one-valued components of $y = \varphi_1(x)$, are commutative, *e. g.* perhaps $F_1[A_1(x)] \equiv A_1[F_1(x)]$. In this case, $y = A_1[F_1(x)]$, occurs twice in the expression $y = \varphi_2(x)$. This requires us to consider the curves as *weighted*, each curve having a positive integer *weight*, corresponding to the number of times that it is to be counted.

One of the most important cases of a multiple-valued real function is the very special case in which $\varphi(x)$ consists of two parts, one of which is simply $y = x$, and the other of which is a real monotonic, increasing, single-valued function, $y = A_1(x)$. We shall obtain in this case,

$$\begin{aligned} \varphi_0(x) &\equiv x \equiv A_0(x), \\ \varphi_1(x) &\equiv \{A_0(x), A_1(x)\}, \\ \varphi_2(x) &\equiv \{A_0(x), 2A_1(x), A_2(x)\}, \\ \varphi_3(x) &\equiv \{A_0(x), 3A_1(x), 3A_2(x), A_3(x)\}, \\ &\cdots\cdots\cdots\cdots \\ \varphi_n(x) &= \{A_0(x), \tbinom{n}{1}A_1(x), \tbinom{n}{2}A_2(x), \cdots, \tbinom{n}{i}A_i(x), \cdots\}, \end{aligned}$$

for n a positive integer. By extending the notion of weight so as to include negative, fractional, and irrational values of the weight, we may define

$\varphi_n(x)$ for all real values of n by the relation

$$\varphi_n(x) \equiv \{A_0(x),\ \tbinom{n}{1}A_1(x),\ \tbinom{n}{2}A_2(x),\ \cdots,\ \tbinom{n}{i}A_n(x),\ \cdots\},$$

the expressions $\binom{n}{i}$ indicating as usual the coefficients in a binomial expansion. It is, furthermore, only by some such extension that the expression $\varphi_{-1}(x)$ can be determined or even defined.

Matrix Interpretation.

16. The iteration of a real function derives a great deal of added significance when regarded as a matrix multiplication. The matrices $\mathbf{A}_n^{(m)}$ which we have hitherto considered have contained but a discrete set of elements. We shall now consider matrices of a more general type, in which the elements that appear in the matrix may form continuous systems. One such type of matrix is suggested by the $\mathbf{A}_1^{(m)}$ already considered. The elements of the $(i + 1)$st row of $\mathbf{A}_1^{(m)}$, are the coefficients of $A_1(x^i)$, where the coefficient of x^j appears in the $(j + 1)$st column, i and j being integers. If now we allow i and j to take on all real positive values, we shall obtain an analogous matrix, in which zeros occur everywhere except on the main diagonal and lines parallel to it starting with the columns at integral units distances from the initial element, and in the first row.

We shall now introduce certain functions $M_n(x, y)$ of two variables, x and y, which functions we shall call *matrices*. The matrix $M_1(x, y)$ associated with a single valued function $y = A_1(x)$, will be the function which vanishes for points not on the locus $y = A_1(x)$, but is unity on the locus. More generally, for a multiple-valued and arbitrarily weighted function $y = \varphi_1(x)$, we shall mean by $M_1(x, y)$ the function whose value at any point (x, y) is equal to the weight of the point, where points not on any component of $y = \varphi_1(x)$ are considered as of weight zero. The matrix product of $M_1(x, y)$ by itself we shall call $M_2(x, y)$ and in general the matrix product of $M_{n-1}(x, y)$ by $M_1(x, y)$ we shall denote by $M_n(x, y)$. The value of $M_2(x, y)$ for any point (x, y) is defined as the algebraic sum of the values of the algebraic product of $M_1(x, t)$ by $M_1(t, y)$, as t varies over the interval a, b, in accordance with the usual notion of matrix product extended in an obvious manner.

If we write $y = \varphi_1(t)$, and $t = \varphi_1(x)$, and give x and y arbitrarily selected values, y_1 and x_1 respectively, the roots of $y_1 = \varphi_1(t)$, $t = (t_1, t_2, t_3, \cdots)$ are in general all distinct from the values of $t = \varphi_1(x_1)$, $t = (t_1', t_2', t_3', \cdots)$. Let however the common values, if any, be denoted by $t_1'', t_2'', t_3'', \cdots$. The weight of the point (t_i'', y_1) we may denote by α_i, $i = 1, 2, \cdots$, and the weight of the point (x_1, t_i''), by β_i, $i = 1, 2, \cdots$. To obtain $y = \varphi_2(x)$, we must eliminate t from $y = \varphi_1(t)$ and $t = \varphi_1(x)$. In particular

the weight of the point (x_1, y_1) will be the sum of the weights given by the points, t_1'', t_2'', t_3'', $\cdots$, or in other words it will be $\Sigma_i \alpha_i \beta_i$. Hence the process of determining φ_2 is much the same as that of finding $M_2(x, y)$; in fact the matrix $M_2(x, y)$ is exactly the matrix M corresponding to $y = \varphi_2(x)$, and, in general, the matrix of $\varphi_n(x)$ is exactly $M_n(x, y)$.

We may now find an entirely new interpretation to one of the formulas already given. Let us denote by $a_{11}z$ a matrix which vanishes except for $x = y$, for which equality, it assumes arbitrary values along this diagonal. By $V_1(z)$ we shall denote a matrix which vanishes along the main diagonal, *i. e.* for $x = y$. By $A_1(z)$ we shall denote the matrix which is the sum of $a_{11}z$ and $V_1(z)$. We shall regard z as a symbolic variable upon which we are operating with the matrices a_{11}, V_1, A_1, etc. The matrix product of the operation $a_{11}z$ by itself we shall denote by $a_{11}{}^2z$ and the inverse of $a_{11}z$ by $(1/a_{11})z$, etc. We shall then have the formula (5) as holding also with this new interpretation, namely,

$$A_n(x) \equiv a_{11}{}^n\left[V_0(z) + \binom{n}{1}\frac{1}{a_{11}}V_1(z) + \binom{n}{2}\frac{1}{a_{11}{}^2}V_2(z) + \cdots \right.$$
$$\left. + \binom{n}{i}\frac{1}{a_{11}{}^i}V_i(z) + \cdots\right],$$

where

$$V_i(z) \equiv A_i(z) - \tbinom{i}{1}a_{11}A_{i-1}(z) + \tbinom{i}{2}a_{11}{}^2A_{i-2}(z) + \cdots \pm a_{11}{}^iA_0(z)$$

and where in the present instance $V_i(z)$ is the ith iterate of $V_1(z)$. It is particularly important as a formula, when $V_1(z)$ represents a curve which is one-valued and monotonic. The weight of $V_1(x)$ may vary from point to point. When a_{11} is equal to unity along the main diagonal and the weight of V_1 is either one or zero, V_1 being the matrix of a one-valued monotonic function, we reobtain the formula given in § 15 for $\varphi_n(x)$.

We have interpreted the iteration of a real function as finding the power of a matrix. It is also possible to interpret finding the power of an ordinary matrix as a case of iteration. For concreteness let us consider the example of the following normal matrix:

$$A_1 \equiv \begin{vmatrix} r_1 & s_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & r_1 & s_{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & r_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r_2 & s_{21} & 0 & 0 \\ 0 & 0 & 0 & 0 & r_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & r_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & r_3 \end{vmatrix}.$$

Here we have for a_{11}

$$a_{11} \equiv \begin{vmatrix} r_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & r_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & r_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & r_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & r_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & r_3 \end{vmatrix},$$

while V_1 is

$$V_1 \equiv \begin{vmatrix} 0 & s_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & s_{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & s_{21} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{vmatrix}.$$

Here V_2 contains but a single element different from zero, and V_3, $V_4 \cdots$, vanish identically. From the formula, we obtain

$$A_n \equiv \begin{vmatrix} r_1{}^n & nr_1{}^{n-1}s_{11} & \dfrac{n(n-1)}{2} r_1{}^{n-2}s_{11}s_{12} & 0 & 0 & 0 & 0 \\ 0 & r_1{}^n & nr_1{}^{n-1}s_{12} & 0 & 0 & 0 & 0 \\ 0 & 0 & r_1{}^n & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r_2{}^n & nr_2{}^{n-1}s_{21} & 0 & 0 \\ 0 & 0 & 0 & 0 & r_2{}^n & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & r_2{}^n & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & r_3{}^n \end{vmatrix},$$

which may be readily verified.*

PART THREE. SOME PROPERTIES OF CERTAIN RATIONAL FUNCTIONS UNDER ITERATION.

A Special Class of Rational Functions.

17. One of the simplest and most interesting classes of functions $y = A_1(x)$ to study in the complex plane, is that of the rational functions.

* Cf. Van Vleck, "One Parameter Projective Groups . . . ," Trans. Am. Math. Soc., 13 (1912), p. 353–386.

Certain rational functions with real coefficients have the property that the iterates of the real function are identical with the iterates considered as real functions. In these cases the positive and negative iterates of a real point are themselves real. It is a class of these functions that we shall now consider.

Let us be given a rational function $A_1(x)$ of degree r in numerator and denominator. For one value of x, $y = A_1(x)$, will have but one value, while for y assigned arbitrarily x must have r values. If these are to be real whenever y is real, which we must require in order that real points are carried into real points by A_{-1}, then the r zeros of $A_1(x)$ must be real and distinct; since if $y = 0$ yields two coincident roots, then $y = c$, where c is properly selected with respect to sign, and different from zero, will have two imaginary roots. Furthermore and for the same reason $y = A_1(x)$ can have no horizontal tangent at a finite point, so that between any two zeros of $y = A_1(x)$, there must be a pole of this function, and this continues to be true when we regard the range cyclically. Thus the poles and zeros alternate. The curve $y = A_1(x)$ must then either increase monotonically between every pair of vertical asymptotes, or else decrease monotonically between every such pair. In either case the expression $\dfrac{dA_1(x)}{dx}$ must have the same sign at all of the zeros of $A_1(x)$. The straight line $y = x$ meets the curve $y = A(x)$ in $r + 1$ points real or imaginary. If the curve decrease monotonically between asymptotes, that is, if $[dA_1(x)]/dx$ is negative at the zeros of $A_1(x)$, then these $r + 1$ points of intersection must all be real and distinct, while for the curve increasing between asymptotes, $r - 1$ at least must be real and distinct, these lying between successive vertical asymptotes, while the two remaining are either conjugate imaginary, real and coincident, or real and distinct.

The sum of the residues of $1/[A_1(x) - x]$, must be zero. We shall suppose that the points of intersection $e_1, e_2, \cdots, e_{r+1}$, of $y = x$ with $y = A_1(x)$ are all finite. We may then obtain a relation among the slopes of the curve at the points e, by use of the residues of $1/[A_1(x) - x]$. The curve $y = A_1(x)$, will be supposed to have r finite poles, and hence it must have a real horizontal asymptote, say, $y = \alpha$. If $1/[A_1(x) - x]$ be expanded in descending powers of x, its first coefficients will be given by using the expression of the form

$$\frac{1}{A_1(x) - x} = -\frac{1}{x - \dfrac{P(x)}{Q(x)}} = \frac{1}{x}\left[\frac{-1}{1 - \dfrac{\dfrac{\alpha}{x} + \dfrac{b_1}{x^2} + \dfrac{b_2}{x^3} + \cdots}{1 + \dfrac{c_1}{x} + \dfrac{c_2}{x^2} + \cdots}}\right] = -\frac{1}{x} - \frac{\alpha}{x^2} + \cdots$$

which determines the residue at infinity of $1/[A_1(x) - x]$ as being equal to unity. If for $dA_1(x)/dx$ we write $A_1'(x)$, the residue at e_i will be $1/[A_1'(e_i)-1]$, hence we have,

$$\sum_{i=1}^{r+1} \frac{1}{1 - A_1'(e_i)} = 1.$$

If $A_1(x)$ decreases between the vertical asymptotes, then $A_1'(x)$ is certainly negative in particular for the $r + 1$ points e_i, $i = 1, 2, \cdots, r + 1$. If it is increasing between the asymptotes, then $A_1'(x)$ is real and greater than unity at each of $r - 1$ points e, one lying between each pair of vertical asymptotes. If the two remaining e's are real and distinct then $A_1'(x)$ will be greater than unity at one of them, and positive but less than unity at the other.

In case $A_1'(x)$ is negative at the points e, there can be no question of a real one-valued function $y = A_n(x)$ which is a continuous function of n, and which leaves one of the e's invariant. For if such an $A_n(x)$ existed its slope at any e, i. e., $A_n'(x)$ would be equal to $[A_1'(x)]^n$ at the same point, but this cannot be real and vary continuously with n and still be negative for $n = 1$. We shall examine now in greater detail the case when a function $y = A_n(x)$ might exist and vary continuously with n, i. e., the case in which $A_1'(x)$ is positive at the real points e. We shall confine ourselves also to the case in which not only $r - 1$, e's are real and distinct but in which the two remaining e's are real and all $r + 1$ are distinct. The cases in which the two remaining e's are conjugate imaginary or coincident may be treated in an analogous manner and present no new features of special interest.

The Iteration of Certain Rational Functions.

18. If we be given a rational function, $y = A_1(x)$, we may, as we have stated before, replace x by a linear fractional transform of x, and y by the same linear fractional transform of y, and the transform of A_1 obtained in this manner has for its iterates the transforms of the iterates of the original A_1. By taking a real linear fractional operation we may arrange matters in such a way that $r - 1$, e's, $e_1, e_2, \cdots, e_{r-1}$, occur each between a pair of vertical asymptotes, while e, and e_{r+1} occur to the right of the right-most asymptote.

We shall suppose therefore that we are given two polynomials $P(x)$ and $Q(x)$ of degree r, and we shall write $A_1(x) \equiv P(x)/Q(x)$. We shall suppose P and Q to be real, and the r zeros of $P(x)$, $z_1, z_2, \cdots, z$, to be real, and the r poles of $A_1(x)$, i. e., the r zeros of $Q(x)$, $p_1, p_2, \cdots, p_r$, to be real while $-\infty < p_1 < z_1 < p_2 < z_2, \cdots, < p_r < z_r < +\infty$. We shall also suppose that $A_1'(z) > 0$ for each point z. There will be $r - 1$ points e_i, $i = 1, 2,$

$\cdots$, $r - 1$, for which $A_1(e_i) = e_i$, and such that $p_i < e_i < p_{i+1}$, while $A_1'(e_i) > 0$. The two remaining roots of $A_1(x) = x$, e_r, and e_{r+1}, we may denote also by f_1 and f_2 respectively. We shall suppose these to be real and distinct, and satisfying the inequality $p_r < f_1 < f_2$. Then $A_1'(f_1) > 1$, and $0 < A_1'(f_2) < 1$.

For n a positive integer greater than one, the function $y = A_n(x)$ has many features analogous to those of $y = A_1(x)$. It is a rational, real function of degree r^n, with r^n real zeros and r^n real poles, and the poles and zeros alternate. The roots of $A_n(x) = x$ are $r^n + 1$ in number of which $r^n - 1$ are real and occur one each between every pair of poles, while the last two are real and to the right of the right-most pole, and coincide indeed with f_1 and f_2. The left-most pole of $A_n(x)$ is to the right of the left-most pole of $A_{n-1}(x)$ but lies also in the interval p_1, p_2. The right-most pole of $A_n(x)$ lies to the right of the right-most pole of $A_{n-1}(x)$ but also in the interval $p_r f_1$. The function $y = A_n(x)$ increases monotonically between every pair of poles, while $A_n'(f_1) > 1$ and $0 < A_n'(f_2) < 1$. If α_1 denotes the asymptotic value of $y = A_1(x)$, then the asymptotic value, α_n, of $A_n(x)$, is such that $\alpha_n < \alpha_1$.

For n positive but nonintegral, the functions $A_n(x)$ which we can determine in the neighborhood of f_1 and of f_2 coincide, in the sense that the series valid at f_1 when analytically extended must be valid at f_2, and must coincide with the series obtained at f_2, for no singularity can occur between these two points. It might be at first supposed that by further continuation we should obtain a function with essential singularities, perhaps, on the real axis, but none the less one-valued in the complex plane. This cannot however be the case if n be allowed to vary continuously, as we shall proceed to show by the use of the Riemann surface for $A_1(x)$. That $y = A_n(x)$ is not one-valued when $n = -1$, and r (the degree) is greater than one, is obvious.

Let us mark in the complex plane of x, the branch points of $y = A_{-1}(x)$, and join these by cuts, to obtain a region R_0, which is simply connected and includes all points of the complex plane not themselves on the cuts. Since f_1 is not itself a branch point we shall suppose that no cut actually passes through f_1, and likewise for f_2. Now the interior of a small circle about f_2 is carried by the transformation $y = A_{-1}(x)$ into the interior of a curve approximately circular and concentric with the given circle, since f_2 is carried into itself, and $A_{-1}'(f_2)$ is real and greater then unity. So that any small region, X, inclosing f_2 is carried into a region Y entirely including X within its interior. Now $A_{-1}(x)$ is defined as one-valued, either finite or infinite, for every x in R_0, so that as X is successively extended Y suffers a like extension and always so as to include X, as long as X is within R_0.

As X comes to coincide with R_0, Y is, of course, no longer a non-overlapping region but coincides with the Riemann surface for $y = A_{-1}(x)$, except for certain cuts, a region which we shall call R_{-1}. We shall define R_{-2} as the multiple-sheeted region that Y becomes when X is taken as R_{-1}, and in general for $X \equiv R_{-n}$ we shall have $Y \equiv R_{-(n+1)}$. On the other hand, for $Y \equiv R_0$, X becomes a region included within R_0 and only partially covering the complex plane. This we shall call R_1. We shall define analogously R_2, R_3, $\cdots$. Now we have a series of regions,

$$\cdots, R_3, R_2, R_1, R_0, R_{-1}, R_{-2}, R_{-3}, \cdots$$

each included in the following, and building up successively an infinitely sheeted surface, which we shall call R.

By the discussion made in the first part of this paper, we may obtain the coefficients of a power series $E(x)$ which represents an analytic function, such that $E[A_{-1}(x)] = (1/a_{11})E(x)$. If we write $t = E(x)$, then the transformation $y = A_{-1}(x)$ corresponds in the t-plane to the transformation $t' = (1/a_{11})t$ where $t' = E[A_{-1}(x)]$. Thus the mapping of R_n upon R_{n-1} has for its transform in the t-plane, a mere magnification leaving the origin invariant. The straight lines through the origin in the t-plane are carried into themselves by the magnification, and are the path curves of the transformation in the t-plane for $t' = (1/a_{11}{}^n)t$. The corresponding path curves in the x-plane, are the path curves for $A_{-n}(x)$, as n varies over real values, and are curves starting at f_2 and passing, in their totality, through every point of R. For $t = E(x)$, as t varies over the t-plane, x varies over R, so that R is the Riemann surface for $t = E(x)$. Now let us take a number x. There will be one point in each sheet of the infinitely-sheeted surface R, having x for its coordinate. Through each of these points there will be a path curve terminating at f_2. To determine $A_n(x)$, for a given real, positive n, we must follow along a path curve from x toward f_2 a distance corresponding to n in the scale that is determined along the path curve by its very nature. The infinite number of points which we shall obtain by starting with the same x but in different sheets, can only be superposed for all choices of x and n when the path curves in the different sheets of R are all superposed. This is not, however, the case unless the rational function be linear fractional, since otherwise $y = A_{-1}(x)$ would carry a single x into but a single y. Hence $A_n(x)$ is not a one-valued function of x as n varies over real positive values.

The function $x = E_{-1}(t)$ is a one-valued meromorphic function of t with infinity as an essential singularity, while $E(x)$ itself is an infinitely many-valued function with R for its Riemann surface.

The same sort of discussion will show that if $A(x)$ be an analytic function

defined with its inverse, all over the plane and $A_1(x)$ is one-valued, then while $A_n(x)$ cannot be one-valued for general values of n, still the inverse E_{-1} of the transforming function E is in general one-valued and meromorphic over the complex plane, provided that there is one point of f used in defining E for which $A_1(f) = f$, and $| A_1'(f) | = A_1'(f) < 1$. We shall not investigate less stringent conditions at this time.

PRINCETON, N. J.,
February, 1915.

THE CAUCHY DEFINITION OF A DEFINITE INTEGRAL.

By D. C. Gillespie.

Cauchy* gave the following definition of the definite integral of a *continuous* function $f(x)$ between the limits (a, b):

$$\int_a^b f(x)dx = \text{Lim}\,[f(a)(x_1-a)+f(x_1)(x_2-x_1)+\cdots+f(x_{m-1})(b-x_{m-1})]. \quad (1)$$

In† justifying his definition he proves that the limit of the sum on the right is the same for all modes of subdivisions of (a, b) in which the limit of the largest sub-interval is zero. He also proves of the integral $\int_a^b f(x)dx$ as thus defined that

$$\int_a^b f(x)dx = \text{Lim}\,[f(\xi_1)(x_1-a)+f(\xi_2)(x_2-x_1)+\cdots+f(\xi_m)(b-x_{m-1})], \quad (2)$$

where $\xi_1\cdots\xi_m$ denote numbers chosen at random in $(a, x_1)\cdots(x_{m-1}, b)$. Riemann‡ takes the limit of the sum (2) as his definition of the definite integral of any function $f(x)$ in the interval (a, b).

A bounded function $f(x)$ will therefore be said to be integrable in the Cauchy sense if the limit on the right in (1) is unique for all modes of subdivision of the interval (a, b) in which the limit of the largest sub-interval is zero; and in the Riemann sense if the like is true of the limit on the right in (2). It is the object of this note to prove that these two definitions are equivalent.

Since the sum (1) is included among the sums (2), if $f(x)$ is integrable in the Riemann sense it is obviously integrable in the Cauchy sense. It is therefore only necessary to prove that if $f(x)$ is not integrable in the Riemann sense it is not integrable in the Cauchy sense.

The necessary§ and sufficient condition that $f(x)$ be integrable in the Riemann sense is that every closed set|| contained in the set of points at which the oscillation** of $f(x)$ is greater than any positive number k that

* Cauchy, Léçons sur le Calcul Infinitesimal, p. 81.

† Cauchy, Oeuvres, Ser. 2, Vol. IV, p. 122–128.

‡ Riemann, Werke, p. 213.

§ Hobson, Theory of Functions of a Real Variable, p. 342.

|| Hobson, loc. cit., p. 64.

** Oscillation-Saltus, Hobson, loc. cit., p. 233.

may be assigned has the content* zero. Hence if $f(x)$ is not integrable in the Riemann sense there is at least one closed set G of content $c(> 0)$ at the points of which the oscillations of $f(x)$ are greater than some constant $k(> 0)$. But it can be shown as follows that if such a set G exists, two systems of subdivisions of (a, b), D_1 and D_2, can be formed for which the limits, granted that they exist, of the sums of the form (1) will differ by $ck/8$; so that $f(x)$ is not integrable in the Cauchy sense.

Let D be a division of the interval (a, b) into m equal subintervals. To prove the existence of D_1 and D_2 we shall consider those intervals of division D which contain points of G either within or at the end points of the intervals. Let x_i, x_{i+1} be the end points of such an interval. If x_i is not a point of G, then there is a first point of G beyond x_i, since G is a closed set. Let us designate this first point of G, provided it lies in the interior of the interval (x_i, x_{i+1}), by p_i. If x_i is a point of G take a point v_i such that $x_i < v_i < x_{i+1}$ and such that $(v_i - x_i) < c/2m$. If v_i is a point of G we shall also designate it by p_i. If v_i is not a point of G, and there are points of G between v_i and x_{i+1}, we shall designate the first of these points beyond v_i by p_i. Apply this process to all the intervals of the division D which contain points of G. Then all the points of G, except possibly b, will lie in a set of intervals, not greater than $2m$ in number, such as (x_i, v_i) and (p_i, x_{i+1}). Each of the intervals (x_i, v_i) is less than $c/2m$; since there are not more than m of them, their sum is less than $c/2$. Therefore the sum of the intervals (p_i, x_{i+1}) must be greater than $c/2$.

We choose now a positive number d_i smaller than the smallest of the three numbers $(p_i - x_i)$, $(x_{i+1} - p_i)$ and $ck/32mM$ (where M is the upper limit of $|f(x)|$ in (a, b)). Then with p_i as center construct an interval of total length $2d_i$. Within this interval about p_i there are two points, which we shall designate by s_i and t_i, such that $f(s_i) - f(t_i) > k/2$.

Let the division D_1 have for end points all the end points of D and in addition the points s_i, and let the division D_2 have for end points all the end points of D and in addition the points t_i. Now the sum (1) due to the division D_1 minus the sum (1) due to the division D_2 is greater than $ck/8$. For the sum (1) due to division D_1 would be made up of terms such as $f(x_i)(s_i - x_i) + f(s_i)(x_{i+1} - s_i)$, together with terms due to those intervals of D in which there is no point p_i. The sum (1) due to division D_2 would be made up of terms such as $f(x_i)(t_i - x_i) + f(t_i)(x_{i+1} - t_i)$, together with terms due to those intervals of D in which there is no point p_i. The difference between the terms of D_1 and D_2 just indicated is

* Hobson, loc. cit., p. 98. Content, as used here, is called upper content by Pierpont, Theory of Functions of Real Variables, p. 352.

$$f(x_i)(s_i - t_i) + f(s_i)(x_{i+1} - p_i + p_i - s_i) - f(t_i)(x_{i+1} - p_i + p_i - t_i)$$
$$= f(x_i)(s_i - t_i) + (f(s_i) - f(t_i))(x_{i+1} - p_i) + f(s_i)(p_i - s_i) - f(t_i)(p_i - t_i).$$

The total difference between the sum (1) due to division D_1 and the sum (1) due to division D_2 would be the sum of all such terms.

$$| \Sigma f(x_i)(s_i - t_i) | < \frac{ck}{16}, \text{ since } | f(x_i) \leqq M \text{ and } |(s_i - t_i)| < 2\delta i < \frac{ck}{16Mm}.$$

$$\Sigma(f(s_i) - f(t_i))(x_{i+1} - p_i) > \frac{ck}{4}, \text{ since } f(s_i) - f(t_i) > \frac{k}{2} \text{ and } \Sigma(x_{i+1} - p_i) > \frac{c}{2}.$$

$$| \Sigma(f(s_i)(p_i - s_i) | < \frac{ck}{32}, \text{ since } | f(s_i)| \leqq M \text{ and } | p_i - s_i | < \delta_i < \frac{ck}{32Mm}.$$

$$| \Sigma(f(t_i)(p_i - t_i) | < \frac{ck}{32}, \text{ since } | f(t_i)| \leqq M \text{ and } | p_i - t_i | < \delta_i < \frac{ck}{32Mm}.$$

Therefore the sum due to D_1 minus the sum due to D_2 is greater than $ck/8$. This difference is independent of m. *Thus as m becomes infinite the limits of the sums due to D_1 and D_2 can not both exist and be equal.*

CORNELL UNIVERSITY.

SURFACES WITH ISOTHERMAL REPRESENTATION OF THEIR LINES OF CURVATURE AS ENVELOPES OF ROLLING.

By Luther Pfahler Eisenhart.

In a former memoir* we have established the existence of transformations of a surface Σ with isothermal representation of its lines of curvature into surfaces Σ_1 of the same kind, such that Σ and a transform Σ_1 envelop a two-parameter family of spheres and the lines of curvature correspond on Σ and Σ_1. Envelopes of spheres possessing the latter property were considered by Ribaucour and are called *transformations of Ribaucour*. It is a property of these transformations that on the locus of the centers of the spheres the curves corresponding to the lines of curvature on Σ and Σ_1 form a conjugate system.

In a recent note Bianchi† developed the idea of the envelope of a plane invariably fixed to a surface S_0 as the latter rolls over an applicable surface S,‡ called the *surface of support*. He called Σ an envelope of rolling and showed that given any non-developable surface Σ the problem of finding pairs of applicable surfaces S_0 and S, such that, as S_0 rolls on S, a plane conveniently fixed with respect to S_0 envelops Σ, requires the solution of a partial differential equation of the second order. In a later note§ Bianchi considered the case where the envelope of rolling Σ is a surface with isothermal representation of its lines of curvature, and found that a transformation of the kind referred to above of Σ into a surface Σ_1 with isothermal representation of its lines of curvature, called by him a *transformation* E_m, gives a solution of the problem. In fact, the locus of the centers of the spheres is the surface of support and the corresponding rolling surface S_0 is found by quadratures.

In the present paper we solve the converse problem: *To show that the transformations E_m are the only transformations of Ribaucour for which the given surface is an envelope of rolling with the locus of the centers of the spheres the surface of support.* In order to establish this result we prove also that the transformations E_m are the only transformations of Ribaucour for

* Transactions of the American Mathematical Society, vol. 9 (1908), pp. 149–177.

† Rendiconti della R. Accademia dei Lincei, vol. 23 (1914), pp. 3–12.

‡ For the meaning of the term rolling as here used see Darboux, Leçons sur la théorie générale des surfaces, Fourth Part, Chapter VI.

§ Rendiconti della R. Accademia dei Lincei, vol. 24 (1915), pp. 366–377.

which the correspondence of the spherical representations of the lines of curvature of the two sheets of the envelope of spheres is conformal.*

2. **Envelope of Rolling.** Let Σ be a surface referred to its lines of curvature; ρ_1, ρ_2 be the principal radii of normal curvature; X, Y, Z the direction-cosines of the normal to Σ; and

$$ds^2 = Edu^2 + Gdv^2 \tag{1}$$

the linear element of Σ. Let M be a point on Σ and M_0 the corresponding point on S, a surface of support. Bianchi has shown that M_0 is on the normal to Σ at M. Hence if x, y, z and x_0, y_0, z_0 are the cartesian coördinates of M and M_0 respectively, we have

$$x_0 = x - RX, \qquad y_0 = y - RY, \qquad z_0 = z - RZ, \tag{2}$$

where R is a function of u and v to be determined. From (2) it is readily found that the linear element of S is

$$ds_0^2 = L_1^2du^2 + L_2^2dv^2 + dR^2,$$

where we have put

$$L_1 = \sqrt{E}\left(1 + \frac{R}{\rho_1}\right), \qquad L_2 = \sqrt{G}\left(1 + \frac{R}{\rho_2}\right). \tag{3}$$

Evidently Σ is one of the sheets of the envelope of the spheres of radius R and center M_0. It is a known property of envelopes of spheres that as the locus of the centers is deformed the two points of contact of each sphere with the envelope retain invariable positions.† Hence when S is applied to S_0, the planes tangent to Σ assume one and the same position in space. If this plane be taken for the xy-plane, the coördinates of S_0 are x, y, R, and consequently its linear element is

$$dx^2 + dy^2 + dR^2.$$

Accordingly as S and S_0 are applicable, we must have

$$L_1^2du^2 + L_2^2dv^2 = dx^2 + dy^2. \tag{4}$$

Since the curvature of the quadratic form on the right is zero, the same must be true of the left hand member. Expressing this condition,‡ we obtain the equation of Bianchi

* In the Rendiconti del Circolo Matematico di Palermo, vol. 39 (1915), pp. 153–176 we established transformations Ω of a surface S into surfaces S_1 such that the common conjugate systems on S and on S_1 have equal tangential invariants and the lines joining corresponding points on S and S_1 form a congruence whose developables meet S and S_1 in these conjugate systems. It was shown that the transformations E_m are the only transformations Ω for which the surfaces S and S_1 envelop a two-parameter family of spheres.

† Bianchi, Lezioni di Geometria Differenziale, vol. 2, p. 88.

‡ Cf. the author's Differential Geometry, p. 155.

$$(5) \qquad \frac{\partial}{\partial u}\left(\frac{\sqrt{G}}{\rho_2}\frac{1}{L_1}\frac{\partial R}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{\sqrt{E}}{\rho_1}\frac{1}{L_2}\frac{\partial R}{\partial v}\right)-\frac{\sqrt{EG}}{\rho_1\rho_2}=0.$$

Conversely, when R satisfies this condition, the reduction of the left hand member of (4) to the form of the right hand member requires at most one quadrature, after which the coördinates of S_0 are known.

3. **Transformations of Ribaucour.** Darboux* has shown that if λ and μ are such that

$$\frac{\partial\lambda}{\partial u}+\rho_1\frac{\partial\mu}{\partial u}=0, \qquad \frac{\partial\lambda}{\partial v}+\rho_2\frac{\partial\mu}{\partial v}=0,$$

the envelope of the spheres of radius λ/μ and center

$$(6) \qquad \xi=x-\frac{\lambda}{\mu}X, \qquad \eta=y-\frac{\lambda}{\mu}Y, \qquad \zeta=z-\frac{\lambda}{\mu}Z,$$

gives the most general transformation of Ribaucour of Σ. We note that the preceding equations are of the same form as the Rodriques equations for Σ, namely

$$\frac{\partial x}{\partial u}+\rho_1\frac{\partial X}{\partial u}=0, \qquad \frac{\partial x}{\partial v}+\rho_2\frac{\partial X}{\partial v}=0.$$

Let X_1, Y_1, Z_1 and X_2, Y_2, Z_2 denote the direction-cosines of the tangents to the curves $v =$ const. and $u =$ const. respectively on Σ. The cartesian coördinates x_1, y_1, z_1 of a transform Σ_1 of Σ are given by equations of the form

$$(7) \qquad x_1=x-\frac{1}{\sigma m}(\mu X+\alpha X_1+\beta X_2),$$

where m is a constant and the other functions satisfy the following completely integrable system of differential equations:†

$$(8) \qquad \begin{aligned}
&\frac{\partial\lambda}{\partial u}=\sqrt{E}\alpha, \qquad \frac{\partial\lambda}{\partial v}=\sqrt{G}\beta,\\
&\frac{\partial\mu}{\partial u}=-\frac{\sqrt{E}}{\rho_1}\alpha, \qquad \frac{\partial\mu}{\partial v}=-\frac{\sqrt{G}}{\rho_2}\beta,\\
&\frac{\partial\alpha}{\partial u}=-\frac{1}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v}\beta+\mu\frac{\sqrt{E}}{\rho_1}+m\sigma(k+\sqrt{E}),\\
&\frac{\partial\alpha}{\partial v}=\frac{\beta}{\sqrt{E}}\frac{\partial\sqrt{G}}{\partial u}, \qquad \frac{\partial\beta}{\partial u}=\frac{\alpha}{\sqrt{G}}\frac{\partial\sqrt{E}}{\partial v},
\end{aligned}$$

* Leçons, vol. 2, p. 383.

† Cf. Eisenhart, Annali di Matematica, vol. 22 (1914), pp. 194–197.

$$\frac{\partial \beta}{\partial v} = -\frac{1}{\sqrt{E}}\frac{\partial \sqrt{G}}{\partial u}\alpha + \mu\frac{\sqrt{G}}{\rho_2} + m\sigma(l + \sqrt{G}),\tag{8}$$

$$\frac{\partial \log \sigma}{\partial u} = k\frac{\alpha}{\lambda}, \qquad \frac{\partial \log \sigma}{\partial v} = l\frac{\beta}{\lambda},$$

$$\frac{\partial k}{\partial v} = l\left[\frac{1}{\sqrt{G}}\frac{\partial \sqrt{E}}{\partial v} - \frac{\beta}{\lambda}(k + \sqrt{E})\right],$$

$$\frac{\partial l}{\partial u} = k\left[\frac{1}{\sqrt{E}}\frac{\partial \sqrt{G}}{\partial u} - \frac{\alpha}{\lambda}(l + \sqrt{G})\right],$$

and the quadratic relation

$$\alpha^2 + \beta^2 + \mu^2 = 2m\lambda\sigma.\tag{9}$$

Conversely every set of functions satisfying equations (8) and (9) determines a transformation of Ribaucour of Σ.

The functions k^2 and l^2 are equal respectively to E_1 and G_1, the first fundamental coefficients of Σ_1. The direction-cosines X_1', Y_1', Z_1'; X_2', Y_2', Z_2' of the tangents to the parametric curves on Σ_1 are given by

$$\frac{\partial x_1}{\partial u} = \sqrt{E_1}X_1', \qquad \frac{\partial x_1}{\partial v} = \sqrt{G_1}X_2',$$

and similar equations in y_1 and z_1. It is necessary to distinguish between two cases, namely

$$(a) \qquad k = -\sqrt{E_1}, \qquad l = \sqrt{G_1},$$

$$(b) \qquad k = \sqrt{E_1}, \qquad l = \sqrt{G_1}.$$

In order that the mutual orientation of the trihedral for Σ_1 shall be the same as for Σ, the direction-cosines of the normal to Σ_1, namely X', Y', Z', are given by equations of the form

$$x_1 - x + \frac{\lambda}{\mu}(X \mp X') = 0,$$

where the upper and lower signs before X' correspond to cases (a) and (b) respectively.

Expressing the condition that the equations of Rodriques for Σ_1 are of the same form as for Σ, we are led to the respective pairs of equations, for the determination of ρ_1' and ρ_2',

$$\begin{aligned}
&\frac{\partial \log \sigma}{\partial u}\left(\pm\frac{1}{\rho_1'} + \frac{\mu}{\lambda}\right) - \frac{\partial}{\partial u}\left(\frac{\mu}{\lambda}\right) = 0,\\
&\frac{\partial \log \sigma}{\partial v}\left(\pm\frac{1}{\rho_2'} + \frac{\mu}{\lambda}\right) - \frac{\partial}{\partial v}\left(\frac{\mu}{\lambda}\right) = 0.
\end{aligned}\tag{10}$$

We consider now case (a). The corresponding equations (10) are reducible by means of (8) to

$$\tag{11} \begin{aligned} \sqrt{E_1}\left(\frac{1}{\rho_1'}+\frac{\mu}{\lambda}\right) &= \sqrt{E}\left(\frac{1}{\rho_1}+\frac{\mu}{\lambda}\right), \\ \sqrt{G_1}\left(\frac{1}{\rho_2'}+\frac{\mu}{\lambda}\right) &= -\sqrt{G}\left(\frac{1}{\rho_2}+\frac{\mu}{\lambda}\right). \end{aligned}$$

4. Solution of the problem. In accordance with the problem as stated we consider now the case where equations (2) and (6) define the same surface. If we put

$$R = \lambda/\mu,$$

it follows from (8) that

$$\frac{\partial R}{\partial u} = \frac{\alpha}{\mu} L_1, \qquad \frac{\partial R}{\partial v} = \frac{\beta}{\mu} L_2.$$

Substituting in equation (5), we reduce the resulting equation by means of (8) and (9) to

$$\frac{\sqrt{G}}{\rho_2}(-\sqrt{E_1}+\sqrt{E}) + \frac{\sqrt{E}}{\rho_1}(\sqrt{G_1}+\sqrt{G}) + 2\frac{\sqrt{EG}}{\rho_1\rho_2}\frac{\lambda}{\mu} = 0,$$

which because of (11) is equivalent to

$$\tag{12} \frac{\sqrt{G}}{\rho_2}\frac{\sqrt{E_1}}{\rho_1'} = \frac{\sqrt{E}}{\rho_1}\frac{\sqrt{G_1}}{\rho_2'}.$$

By similar processes we find for case (b) the equation

$$\tag{12'} \frac{\sqrt{G}}{\rho_2}\frac{\sqrt{E_1}}{\rho_1'} + \frac{\sqrt{E}}{\rho_1}\frac{\sqrt{G_1}}{\rho_2'} = 0.$$

If the linear elements of the spherical representations of Σ and Σ_1 be written

$$ds'^2 = e\,du^2 + g\,dv^2, \qquad ds_1'^2 = e_1du^2 + g_1dv^2,$$

then*

$$\sqrt{e} = \frac{\sqrt{E}}{\rho_1}, \qquad \sqrt{g} = \frac{\sqrt{G}}{\rho_2}, \qquad \sqrt{e_1} = \frac{\sqrt{E_1}}{\rho_1'}, \qquad \sqrt{g_1} = \frac{\sqrt{G_1}}{\rho_2'}.$$

Also we have†

$$\tag{13} \frac{\partial\sqrt{e}}{\partial v} = \frac{1}{\rho_2}\frac{\partial\sqrt{E}}{\partial v}, \qquad \frac{\partial\sqrt{g}}{\partial u} = \frac{1}{\rho_1}\frac{\partial\sqrt{G}}{\partial u}.$$

Consequently equations (12) and (12′) give

---

* Cf. the author's Differential Geometry, p. 200.

† Ibid., p. 157.

$$(14) \qquad \frac{e}{e_1} = \frac{g}{g_1},$$

from which it follows that

A necessary and sufficient condition that the locus of the centers of the spheres of a transformation of Ribaucour of a surface Σ be the surface of support with Σ the corresponding envelope of rolling is that the correspondence between the spherical representations of the lines of curvature of the two sheets of the envelope of spheres be conformal.

5. **Transformations** E_m. Now we shall show that transformations E_m are characterized by the property mentioned in the above theorem.

We replace equation (12) by

$$\frac{\sqrt{E_1}}{\rho_1'} = \rho \frac{\sqrt{E}}{\rho_1}, \qquad \frac{\sqrt{G_1}}{\rho_2'} = \rho \frac{\sqrt{G}}{\rho_2},$$

where the factor of proportionality ρ is to be determined. Now equations (11) may be replaced by

$$(15) \qquad \sqrt{E_1} - \sqrt{E} = \sqrt{e}\frac{\lambda}{\mu}(1 - \rho), \qquad \sqrt{G_1} + \sqrt{G} = -\sqrt{g}\frac{\lambda}{\mu}(1 + \rho).$$

If these values of $\sqrt{E_1}$ and $\sqrt{G_1}$ be substituted in the last two of equations (8), we find that ρ must satisfy the equations

$$(16) \qquad \begin{aligned} \frac{\partial}{\partial v}\log \rho &= -\frac{\partial}{\partial v}\log e - \frac{\beta}{\mu}(1 - \rho)g, \\ \frac{\partial}{\partial u}\log \rho &= -\frac{\partial}{\partial u}\log g - \frac{\alpha}{\mu}(1 + \rho)e. \end{aligned}$$

Expressing the condition of integrability of these equations, we get

$$\frac{\partial^2}{\partial u \partial v}\log\frac{e}{g} = 0.$$

Hence

$$\frac{e}{g} = \frac{U}{V},$$

where U and V are functions of u and v respectively, that is, the spherical representation of the lines of curvature of Σ is isothermal; and in consequence of (14) the same is true of Σ_1.

When similar processes are applied to (12′), we get the same equations (16).

By a suitable choice of the parameters we can reduce U and V to unity and then we introduce functions θ and φ by the equations

$$(17) \qquad \sqrt{e} = \sqrt{g} = e^{-\theta}, \qquad \rho = e^{2\theta}\mu/\varphi,$$

where the e's of the right hand members are the Naperian base. Now equations (16) reduce to

$$\frac{\partial \varphi}{\partial u} = \alpha e^{\theta}, \qquad \frac{\partial \varphi}{\partial v} = -\beta e^{\theta}, \tag{18}$$

which are readily found to be consistent.

From (15) and (17) we have

$$\sqrt{E_1} = \sqrt{E} + \frac{\lambda}{\mu}\left(e^{-\theta} - e^{\theta}\frac{\mu}{\varphi}\right), \qquad \sqrt{G_1} = -\sqrt{G} - \frac{\lambda}{\mu}\left(e^{-\theta} + e^{\theta}\frac{\mu}{\varphi}\right). \tag{19}$$

Substituting these values in the expressions (8) for $\partial \log \sigma/\partial u$ and $\partial \log \sigma/\partial v$, we have

$$\frac{\partial \log \sigma}{\partial u} = -\frac{\partial \log \lambda}{\partial u} + \frac{\partial \log \mu\varphi}{\partial u}, \qquad \frac{\partial \log \sigma}{\partial v} = -\frac{\partial \log \lambda}{\partial v} + \frac{\partial \log \mu\varphi}{\partial v}.$$

Hence to within a constant factor we have by integration

$$\sigma = \mu\varphi/\lambda. \tag{20}$$

This constant will be taken equal to unity, as the constant m appears in (8) and (9) as a multiplier of σ.

Now the first eight of equations (8) and (18) are reducible to

$$\begin{gathered}
\frac{\partial \lambda}{\partial u} = \sqrt{E}\alpha, \qquad \frac{\partial \lambda}{\partial v} = \sqrt{G}\beta, \\
\frac{\partial \mu}{\partial u} = -e^{-\theta}\alpha, \qquad \frac{\partial \mu}{\partial v} = -e^{-\theta}\beta, \\
\frac{\partial \alpha}{\partial u} = \frac{\partial \theta}{\partial v}\beta + e^{-\theta}\mu - m(e^{-\theta}\varphi - e^{\theta}\mu), \\
\frac{\partial \alpha}{\partial v} = -\frac{\partial \theta}{\partial u}\beta, \qquad \frac{\partial \beta}{\partial u} = -\frac{\partial \theta}{\partial v}\alpha, \\
\frac{\partial \beta}{\partial v} = \frac{\partial \theta}{\partial u}\alpha + e^{-\theta}\mu - m(e^{-\theta}\varphi + e^{\theta}\mu), \\
\frac{\partial \varphi}{\partial u} = \alpha e^{\theta}, \qquad \frac{\partial \varphi}{\partial v} = -\beta e^{\theta}.
\end{gathered} \tag{21}$$

These equations are obtained in case (b) also, and the only change necessary is that of the sign of $\sqrt{E_1}$ in (19).

Equations (21) are equivalent to the equations of a transformation E_m derived by Bianchi.* Hence we have

* L. c., p. 371.

A necessary and sufficient condition that the correspondence between the spherical representations of the lines of curvature of two surfaces in the relation of a transformation of Ribaucour be conformal is that the spherical representations be isothermal, in which case the transformation is an E_m.

Combining these two theorems, we have the desired result:

Transformations E_m *are the only transformations of Ribaucour for which the given surface is an envelope of rolling with the locus of the centers of the spheres for the surface of support.*

PRINCETON UNIVERSITY.

A THEOREM CONCERNING REAL FUNCTIONS.

By K. P. Williams.

The purpose of this note is to state a theorem that can be made to apply to any function not a constant defined on a linear interval, and which has a corollary one of the most important theorems concerning continuous functions. As the demonstration is very simple it is thought that the theorem can well be used in an introductory course on the real variable. While in such courses the Dirichlet definition of a function may be given, no general properties are usually obtained until some rather strong hypothesis, such as continuity, is imposed upon the function.

Let $f(x)$ be a function defined throughout the closed interval (a, b), and suppose it is not constantly of the same sign, and not everywhere zero. Then there exists at least one point such that in no vicinity of the point is the function constantly of the same sign.

Suppose the theorem not true. Then about each point of (a, b) we can construct an interval throughout which the function is of the same sign or zero. These intervals will overlap, so that by the Heine-Borel theorem a finite number of them covers (a, b). As these also overlap it follows that the function must be of the same sign throughout (a, b), or else everywhere zero, which is contrary to hypothesis. Thus the theorem must be true.

If $f(x)$ is continuous and both positive and negative, it follows at once that it is somewhere zero.

It is apparent that the theorem can be modified so as to generalize the theorem that a continuous function actually assumes all values between any two of its values. Let $f(x)$ be any function not constant, and M and m its upper and lower bounds, respectively. Then if l is any number between m and M there must be at least one point in no vicinity of which $l - f(x)$ is of the same sign.

The theorem can be generalized so as to apply to a function defined on a very general range. In the form that follows the terms used are those employed by Fréchet, "Sur Quelques Points du Calcul Fonctionel," Rendiconti del Circolo Matematico di Palermo, T. XXII, 1–42.

Let U be an operation defined for each element of an extremal class E, which is itself a sub-class of elements of a normal class. Then either there exists an element h of E such that on no sub-class of E to which h is properly interior is U constantly of the same sign, or else E can be broken up into a

finite number of closed sets, having no elements in common, and on each of which U is constantly of the same sign or constantly zero.

Suppose there exists no point such as h. Then each element a of E gives rise to a sub-class E_a, to which a is properly interior and on which U is constantly of the same sign or constantly zero. From the generalized Heine-Borel theorem it follows that there is a finite number of these classes, call them $E_{a_1}, E_{a_2}, \cdots, E_{a_m}$, such that every element of E is interior to one of them. In case two of these classes have an element in common, they can be combined into a single class on which the function is constantly of the same sign or constantly zero. Suppose next that an element d of E_{a_i} is a limit element for elements of E_{a_j}. Since d is properly interior to some one of the classes we see that U is of the same sign or zero throughout the combined class $E_{a_i} + E_{a_j}$. It follows from this that the m sets $E_{a_1} E_{a_2} \cdots$, E_{a_m} can be replaced by n sets $(n < m)$, call them $E_1, E_2, \cdots, E_n$, which are all closed, have no elements in common, and on each of which U is constantly of the same sign or constantly zero. Thus the theorem is proved.

UNIVERSITY OF INDIANA.

NOTE ON AN OPERATION OF THE THIRD GRADE.

By Albert A. Bennett.

If a and b be numbers, the operation of adding b to a has been called an operation of the first grade (*Stufe*), that of multiplying a by b, an operation of the second grade, and that of raising a to the power b, one of the third grade. With this definition an operation of the third grade bears a relation to one of the second grade different from that existing between the first two grades. Now $a \times b$ may be written, when dealing with ordinary numbers, as $e^{\log a + \log b}$. If in this expression we replace addition by multiplication we shall have $e^{(\log a)(\log b)} = e^{e^{\log \log a + \log \log b}}$, which we may define as $a{}^0_2b$. We may similarly define $a{}^0_3b$ as $e^{\log a {}^0_2 \log b}$, and in general for n an integer, $a{}^0_nb$ as $e^{\log a {}^0_{n-1} \log b}$, where in particular $a{}^0_0b = a + b$ and $a{}^0_1b = a \times b$. An operation, ${}^0_2$, exists for every modular field and for each choice of a primitive root of the field, but no analogous operation, ${}^0_3$, can be defined for modular fields, since the set of indices obtained as exponents of the primitive root, do not themselves form a field having a primitive root.

The question arises as to whether for the ordinary real or complex number systems we are able to interpolate in the series of operations, ${}^0_n$, operations, such that ${}^0_n$ may be defined continuously for all real or complex values of n. We shall be able to do so when once we have obtained the iteration of e^x for all values of n, the index of iteration, or as we may write it $E_n(x)$, where $E_1(x) \equiv e^x$ and $E_2(x) = e^{e^x}$, etc.

We can obtain a definition of $E_n(x)$, even when we restrict ourselves to real numbers, by finding a function $f(x)$ such that $e^{f(x)} = f(x+1)$. With such an $f(x)$, we may write $E_n(x) \equiv f[n + f^{-1}(x)]$. We may take $f(0) = -\infty$, so that $f(1) = 0$, while between 0 and 1, $f(x)$ is defined as a real one-valued function. It may, however, have finite or infinite discontinuities of an arbitrary type between 0 and 1, but it will in any case be completely defined for all positive values of x when once defined for $0 \leq x \leq 1$. It is natural to seek to determine a particular $f(x)$ such that $y = f(x)$ is as regular as possible as x approaches $+\infty$. Since $f(+\infty) = +\infty$, we shall use $g(x) \equiv \dfrac{1}{f(x)}$. A solution $g(x)$ may be found which is asymptotic to $0 + \dfrac{0}{x} + \dfrac{0}{x^2} + \cdots + \dfrac{0}{x^n} + \cdots$. It might be at first supposed that but one solution $g(x)$ of $e^{1/g(x)} = 1/g(x+1)$ could be found for which $g(1) = \infty$

and which is asymptotic to $0 + \frac{0}{x} + \cdots + \frac{0}{x^n} + \cdots$, but such is not the case. For example, if $f(x)$ be defined between 0 and 1 as consisting of a finite number of arbitrary portions of arbitrarily selected meromorphic functions, and meromorphic at each point of discontinuity both from the right and from the left, then $g(x) \equiv 1/f(x)$ will of necessity be asymptotic to the above series. Thus the asymptotic behavior of $g(x)$ does not serve to identify it.

Probably the simplest method of obtaining a function $f(x)$ satisfying $e^{f(x)} = f(x+1)$ which shall be continuous together with all of its derivatives for positive real values of x is to resort to the complex plane, and apply the usual methods for finding the iteration $E_n(x)$.* Thus we seek first a number c such that $e^c = c$. Any such point $c = u + iv$, will be given by the intersections of the curves $v^2 = (e^u)^2 - u^2$ and $u = v \cot v$. These intersections are isolated, infinite in number, and are situated symmetrically with respect to the real axis. If c_1 is the root, above the real axis, of smallest absolute value, the series expansion obtained for $E_n(x)$ at c_1, by the usual method will determine the same function on the positive real axis as when we start with $\bar{c}_1$, the conjugate of c_1. In a strip lying to the positive side of the pure imaginary axis, including the real axis, c_1 and $\bar{c}_1$, but not including any other root c, $E_n(x)$ thus defined will be one-valued, analytic, and will carry real points into real points, regarded as a function of x, for any positive real value of n.

It might have been supposed that since $f(x)$ is determinate for $x = 1, 2, \cdots, n, \cdots$, we might apply Newton's interpolation formula, and have obtained an analytic function $f(x)$ as desired, although confining ourselves to real numbers. The function obtained by the interpolation formula is, however, readily seen not to satisfy $e^{f(x)} = f(x+1)$, since, for example, it will be finite for $x = 0$. When, however, we use the point c_1, the function $E_n(x)$ obtained is exactly the one obtained by a slight modification of Newton's formula. We have merely to expand by Newton's formula, $E_n(x)/a^n$, where the first n is the index of iteration, and the second denotes raising to a power, and where $a = (e^x)' \mid_{x=c_1}$, *i. e.*, $a = c_1$. Thus, with the usual notation, if $y_n(x) \equiv E_n(x)/c_1{}^n$, for integer values of n, then

$$E^n(x) = c_1{}^n\left[y_0(x) + \binom{n}{1}\Delta y_0(x) + \binom{n}{2}\Delta^2 y_0(x) + \cdots + \binom{n}{r}\Delta^r y_0(x) + \cdots\right].$$

PRINCETON, N. J.,
September, 1915.

* Cf. article by author, *Annals of Mathematics*, Vol. 17 (1915), p. 31, ff. The expansion given below is identical with eq. (5), p. 34.

DETERMINATION OF ALL TRIPLY ORTHOGONAL SYSTEMS CONTAINING A FAMILY OF MINIMAL SURFACES.*

By T. H. Gronwall.

1. Introduction.

Two systems with the property indicated in the title are known:

First, the general triply orthogonal system containing a family of planes, and generated by tracing any orthogonal system of curves in a plane and letting this plane roll on a developable surface.†

Second, the system containing a family of minimal surfaces of revolution, that is, catenoids, the axis of revolution being the same for each individual surface. This system appears as the last case in the present investigation (eq. 58 and 59). It may be observed that this case falls under the first, since every triply orthogonal system containing a family of surfaces of revolution also contains a family of planes passing through the common axis of revolution. To these may be added at once the systèms containing an arbitrary family of imaginary minimal cylinders represented by the equations in art. 189 of Darboux's Théorie des surfaces, vol. I, 2d ed., p. 341. This case, and that of the planes, exhaust the cases where the minimal surfaces cannot be represented by formulas (1) in § 2.

In the present paper, it is shown that besides the cases indicated there exists only one family of minimal surfaces belonging to a triply orthogonal system, namely the imaginary quartic surfaces

$$\frac{1}{432k(\rho)^2}(x+iy)^4+(x+iy)(x-iy+l(\rho))+z^2=0,$$

where $k(\rho)$ and $l(\rho)$ are arbitrary functions of the parameter ρ. In § 2, the differential equation of the problem—eq. (6)—is derived, and since the direct integration of this equation would present great difficulties, a preliminary function theoretic examination of the singular points of the equation is given in § 3, the result permitting the restriction of the solutions, a priori, to a small number of types, falling into two classes, to each of which one of the following two paragraphs is devoted. In these paragraphs function theoretic methods are used freely and enable us to keep the necessary algebraic calculations within very moderate bounds.

* Read, in part, before the American Mathematical Society, April 1914.

† Darboux, Leçons sur les systèmes orthogonaux et les coordonnées curvilignes, 2d ed., Paris, 1910, chapter 2.

2. The differential equation of the problem.

Disregarding, from the outset, the case of a plane, or the imaginary cylinders mentioned in the introduction, any other minimal surface may be represented by the Weierstrass formulas*

$$
(1)\qquad
\begin{aligned}
x + iy &= -u^2f'' + 2uf' - 2f + g'',\\
x - iy &= f'' - v^2g'' + 2vg' - 2g,\\
z &= uf'' - f' + vg'' - g',
\end{aligned}
$$

where f and g are analytic functions of u and v respectively. For a *real* minimal surface, u and v are conjugate complex variables, and f''' and g''' are conjugate functions. From (1) we obtain

$$
(2)\qquad
\begin{aligned}
&\frac{\partial}{\partial u}(x + iy) = -u^2f''', \qquad \frac{\partial}{\partial u}(x - iy) = f''', \qquad \frac{\partial z}{\partial u} = uf''',\\
&\frac{\partial}{\partial v}(x + iy) = g''', \qquad \frac{\partial}{\partial v}(x - iy) = -v^2g''', \qquad \frac{\partial z}{\partial v} = vg''',
\end{aligned}
$$

and

$$
(3)\qquad
\begin{aligned}
\frac{\partial}{\partial u}(x^2 + y^2 + z^2) &= \frac{\partial}{\partial u}[(x + iy)(x - iy) + z^2]\\
&= [-2f + (1 + uv)^2g'' - 2u(1 + uv)g' + 2u^2g]f''',\\
\frac{\partial}{\partial v}(x^2 + y^2 + z^2) &= [-2g + (1 + uv)^2f'' - 2v(1 + uv)f' + 2v^2f]g'''.
\end{aligned}
$$

We now form the partial differential equation defining the conjugate system formed by the lines of curvature on the surface (1), that is, an equation of the form

$$
A\frac{\partial^2\theta}{\partial u^2} + B\frac{\partial^2\theta}{\partial u\partial v} + C\frac{\partial^2\theta}{\partial v^2} + D\frac{\partial\theta}{\partial u} + E\frac{\partial\theta}{\partial v} = 0
$$

(where $A = A(u, v)$, etc.), admitting the particular solutions $x + iy$, $x - iy$, z and $x^2 + y^2 + z^2$. This equation is found to be

$$
(4)\qquad \frac{\partial}{\partial u}\left(\frac{1}{(1 + uv)^2f'''}\cdot\frac{\partial\theta}{\partial u}\right) = \frac{\partial}{\partial v}\left(\frac{1}{(1 + uv)^2g'''}\cdot\frac{\partial\theta}{\partial v}\right),
$$

as may be verified by (2) and (3) in the following manner:

First, $x + iy$ is a solution, since

$$
\frac{\partial}{\partial u}\frac{-u^2}{(1 + uv)^2} = \frac{\partial^2}{\partial u\partial v}\frac{u}{1 + uv} = \frac{\partial}{\partial v}\frac{1}{(1 + uv)^2};
$$

* Darboux, Théorie des surfaces I, 2d ed., p. 340.

second, $x - iy$ is a solution, since

$$\frac{\partial}{\partial u}\frac{1}{(1+uv)^2} = \frac{\partial^2}{\partial u \partial v}\frac{v}{1+uv} = \frac{\partial}{\partial v}\frac{-v^2}{(1+uv)^2};$$

third, z is a solution, since

$$\frac{\partial}{\partial u}\frac{u}{(1+uv)^2} = \frac{\partial^2}{\partial u \partial v}\frac{-1}{1+uv} = \frac{\partial}{\partial v}\frac{v}{(1+uv)^2},$$

and finally, $x^2 + y^2 + z^2$ is a solution, since

$$\begin{aligned}\frac{\partial}{\partial u}\left(-\frac{2f}{(1+uv)^2} + g'' - \frac{2u}{1+uv}g' + \frac{2u^2}{(1+uv)^2}g\right)& \\ = -\frac{\partial^2}{\partial u \partial v}\frac{2(vf+ug)}{1+uv}& \\ = \frac{\partial}{\partial v}\left(f'' - \frac{2v}{1+uv}f' + \frac{2v^2}{(1+uv)^2}f - \frac{2g}{(1+uv)^2}\right)&.\end{aligned}$$

If in (1) we let f and g depend also on a parameter ρ: $f = f(u, \rho)$ and $g = g(v, \rho)$, we obtain a family of minimal surfaces. Let $Hd\rho$ denote the distance, at the point x, y, z, between the two surfaces corresponding to the parameters ρ and $\rho + d\rho$; then, when x, y and ρ are taken as independent variables, we obviously have

$$H = \frac{\dfrac{\partial z}{\partial \rho}}{\sqrt{1 + \left(\dfrac{\partial z}{\partial x}\right)^2 + \left(\dfrac{\partial z}{\partial y}\right)^2}},$$

and taking u, v and ρ as independent variables, this expression is transformed into

$$H = \frac{\dfrac{\partial(x, y, z)}{\partial(u, v, \rho)}}{\sqrt{\left(\dfrac{\partial(x, y)}{\partial(u, v)}\right)^2 + \left(\dfrac{\partial(y, z)}{\partial(u, v)}\right)^2 + \left(\dfrac{\partial(z, x)}{\partial(u, v)}\right)^2}}.$$

Calculating H for the minimal surfaces (1), we readily obtain

$$(5)\quad H = \frac{\partial^2 f(u, \rho)}{\partial u \partial \rho} + \frac{\partial^2 g(v, \rho)}{\partial v \partial \rho} - \frac{2v}{1+uv}\frac{\partial f(u, \rho)}{\partial \rho} - \frac{2u}{1+uv}\frac{\partial g(v, \rho)}{\partial \rho}.$$

*In order that the family (1) of minimal surfaces belong to a triply orthogonal system, it is necessary and sufficient that the expression (5) for H shall be a solution of the differential equation (4).**

* Darboux, Leçons sur les systèmes orthogonaux et les coordonnées curvilignes, 2d ed., p. 76.

Since (4) and (5) remain unchanged when u is replaced by v and f by g, the substitution of (5) in (4) gives an equation which, upon multiplication by $(1 + uv)^5$, takes the form

$$\{f, g, u, v\} = \{g, f, v, u\}, \tag{6}$$

where

$$\begin{aligned} \{f, g, u, v\} = {} & (1 + uv)^3 \frac{\partial}{\partial u}\left(\xi \frac{\partial^3 f}{\partial u^2 \partial \rho}\right) - 2v(1 + uv)^2 \xi \frac{\partial^3 f}{\partial u^2 \partial \rho} \\ & - 2v(1 + uv)^2 \frac{\partial}{\partial u}\left(\xi \frac{\partial^2 f}{\partial u \partial \rho}\right) + 6v^2(1 + uv)\xi \frac{\partial^2 f}{\partial u \partial \rho} \\ & + 2v^2(1 + uv) \frac{\partial}{\partial u}\left(\xi \frac{\partial f}{\partial \rho}\right) - 8v^3 \xi \frac{\partial f}{\partial \rho} \\ & - 2(1 + uv) \frac{\partial g}{\partial \rho} \frac{\partial \xi}{\partial u} + 8v \frac{\partial g}{\partial \rho} \xi, \end{aligned} \tag{7}$$

and

$$\xi = \xi(u, \rho) = \frac{1}{\dfrac{\partial^3 f(u, \rho)}{\partial u^3}}, \qquad \eta = \eta(v, \rho) = \frac{1}{\dfrac{\partial^3 g(v, \rho)}{\partial v^3}}. \tag{8}$$

Equation (6) is the differential equation of our problem.

3. Preliminary discussion of equation (6) and classification of possible solutions.

We begin by proving that, if f, g is a solution of (6), and we give ρ an arbitrary but constant value, either the pair of functions ξ and $\partial f/\partial \rho$ or the pair η and $\partial g/\partial \rho$ (or both pairs) has only isolated singular points. The method used will also give a first classification of the solutions.

To this purpose, we differentiate (6) four times in respect to u and four times in respect to v; after the necessary algebraic reductions, we obtain

$$\begin{aligned} & \frac{\partial^5 \xi}{\partial u^5} \cdot \frac{\partial^5 g}{\partial v^4 \partial \rho} + u \frac{\partial^5 \xi}{\partial u^5} \cdot \left(v \frac{\partial^5 g}{\partial v^4 \partial \rho} + 4 \frac{\partial^4 g}{\partial v^3 \partial \rho}\right) \\ & \qquad = \frac{\partial^5 f}{\partial u^4 \partial \rho} \cdot \frac{\partial^5 \eta}{\partial v^5} + \left(u \frac{\partial^5 f}{\partial u^4 \partial \rho} + 4 \frac{\partial^4 f}{\partial u^3 \partial \rho}\right) \cdot v \frac{\partial^5 \eta}{\partial v^5}. \end{aligned} \tag{9}$$

Writing

$$\begin{aligned} \varphi_1(u) &= \frac{\partial^5 \xi}{\partial u^5}, & \varphi_2(u) &= \frac{\partial^5 f}{\partial u^4 \partial \rho}, & \varphi_3(u) &= u \frac{\partial^5 f}{\partial u^4 \partial \rho} + 4 \frac{\partial^4 f}{\partial u^3 \partial \rho}, \\ \psi_1(v) &= \frac{\partial^5 \eta}{\partial v^5}, & \psi_2(v) &= \frac{\partial^5 g}{\partial v^4 \partial \rho}, & \psi_3(v) &= v \frac{\partial^5 g}{\partial v^4 \partial \rho} + 4 \frac{\partial^4 g}{\partial v^3 \partial \rho}, \end{aligned} \tag{10}$$

(9) takes the form

$$\varphi_1(u)\psi_2(v) + u\varphi_1(u)\psi_3(v) = \psi_1(v)\varphi_2(u) + v\psi_1(v)\varphi_3(u), \tag{11}$$

and here we may distinguish three cases:

Case I: $\varphi_1(u) = 0$, $\psi_1(v) \neq 0$.

Then (11) reduces to $\varphi_2(u) + v\varphi_3(u) = 0$ whence $\varphi_2(u) = \varphi_3(u) = 0$. $\varphi_1(u) = 0$ gives $\xi = P_4(u, \rho)$, where P_4 is a polynomial of the fourth degree in u; and $\varphi_2(u) = \varphi_3(u) = 0$ gives $\partial^4 f/\partial u^3 \partial \rho = 0$ or, by (8), $\partial\xi/\partial\rho = 0$, so that $P_4 = P_4(u)$ has its coefficients independent of ρ, and the same equation gives $\partial f/\partial \rho = P_2(u, \rho)$, a polynomial of the second degree in u. Therefore, finally,

$$\frac{\partial f}{\partial \rho} = A_0'(\rho) + A_1'(\rho)u + A_2'(\rho)u^2, \qquad \frac{\partial^3 f}{\partial u^3} = \frac{1}{P_4(u)}, \tag{12}$$

so that the only singularities of ξ and $\partial f/\partial\rho$ are at infinity. The case $\psi_1(v) = 0$, $\varphi_1(u) \neq 0$ evidently reduces to case I by interchanging u and v, f and g.

Case II: $\varphi_1(u) = \psi_1(v) = 0$.

Equations (10) and (8) give immediately $\xi = P_4(u, \rho)$, $\eta = Q_4(v, \rho)$ or

$$\frac{\partial^3 f}{\partial u^3} = \frac{1}{P_4(u, \rho)}, \qquad \frac{\partial^3 g}{\partial v^3} = \frac{1}{Q_4(v, \rho)}, \tag{13}$$

where P_4 and Q_4 are polynomials of the fourth degree in u and v respectively; the coefficients, however, are not necessarily independent of ρ. The only singularities of $\partial f/\partial\rho$ and $\partial g/\partial\rho$ at finite distance are the zeros of P_4 and Q_4.

Case III: $\varphi_1(u) \neq 0$, $\psi_1(v) \neq 0$.

Writing (11) in the form

$$\frac{\psi_2(v)}{\psi_1(v)} + u\frac{\psi_3(v)}{\psi_1(v)} = \frac{\varphi_2(u)}{\varphi_1(u)} + v\frac{\varphi_3(u)}{\varphi_1(u)},$$

differentiating once in respect to v, and then giving v a constant value, it is seen that $\varphi_3(u)/\varphi_1(u)$ must be a linear function of u. Substituting this in our equation, we find the same to be true for $\varphi_2(u)/\varphi_1(u)$, so that

$$\begin{aligned} \varphi_2(u) &= (a + bu)\varphi_1(u), \\ \varphi_3(u) &= (c + du)\varphi_1(u), \end{aligned} \tag{14}$$

where the constants a, b, c and d will of course depend on the parameter ρ. Substituting (14) in (11), we have

$$\begin{aligned} \psi_2(v) &= (a + cv)\psi_1(v), \\ \psi_3(v) &= (b + dv)\psi_1(v). \end{aligned} \tag{15}$$

By means of (10), equations (14) may be written

$$\frac{\partial^5 f}{\partial u^4 \partial \rho} = (a + bu) \frac{\partial^5 \xi}{\partial u^5},$$

$$u \frac{\partial^5 f}{\partial u^4 \partial \rho} + 4 \frac{\partial^4 f}{\partial u^3 \partial \rho} = (c + du) \frac{\partial^5 \xi}{\partial u^5},$$

or

$$\frac{\partial^5 f}{\partial u^4 \partial \rho} = \frac{\partial^4}{\partial u^4}\left[(a + bu) \frac{\partial \xi}{\partial u} \right] - 4b \frac{\partial^4 \xi}{\partial u^4},$$

$$\frac{\partial^4}{\partial u^4}\left(u \frac{\partial f}{\partial \rho} \right) = \frac{\partial^4}{\partial u^4}\left[(c + du) \frac{\partial \xi}{\partial u} \right] - 4d \frac{\partial^4 \xi}{\partial u^4},$$

whence, integrating four times in respect to u,

$$(16) \qquad \begin{aligned} \frac{\partial f}{\partial \rho} &= (a + bu) \frac{\partial \xi}{\partial u} - 4b\xi + A_1 u^3 + B_1 u^2 + C_1 u + D_1, \\ u \frac{\partial f}{\partial \rho} &= (c + du) \frac{\partial \xi}{\partial u} - 4d\xi + A_2 u^3 + B_2 u^2 + C_2 u + D_2. \end{aligned}$$

The elimination of $\partial f / \partial \rho$ between these two equations gives

$$(17) \qquad \begin{aligned} [bu^2 + (a - d)u - c] \frac{\partial \xi}{\partial u} - 4(bu - d)\xi + A_1 u^4 + (B_1 - A_2)u^3 \\ + (C_1 - B_2)u^2 + (D_1 - C_2)u - D_2 = 0. \end{aligned}$$

Similarly, we obtain from (15)

$$(18) \qquad \begin{aligned} \frac{\partial g}{\partial \rho} &= (a + cv) \frac{\partial \eta}{\partial v} - 4c\eta + A_3 v^3 + B_3 v^2 + C_3 v + D_3, \\ v \frac{\partial g}{\partial \rho} &= (b + dv) \frac{\partial \eta}{\partial v} - 4d\eta + A_4 v^3 + B_4 v^2 + C_4 v + D_4, \end{aligned}$$

$$(19) \qquad \begin{aligned} [cv^2 + (a - d)v - b] \frac{\partial \eta}{\partial v} - 4(cv - d)\eta + A_3 v^4 + (B_3 - A_4)v^3 \\ + (C_3 - B_4)v^2 + (D_3 - C_4)v - D_4 = 0. \end{aligned}$$

From the first equation (16) it is clear that the only singularities of $\partial f / \partial \rho$ at finite distance are those of ξ, and these, according to the linear differential equation (17) defining ξ, cannot occur elsewhere than at the zeros of $bu^2 + (a - d)u - c$, or, should this expression vanish identically, at the zero of $bu - d$. The same argument applies to η and $\partial g / \partial \rho$. The argument fails, however, when also $bu - d$ vanishes identically; but then

$$a = b = c = d = 0,$$

and from (17) and (19) we obtain $A_1 = A_3 = 0$, and (16) and (18) give

$$\frac{\partial f}{\partial \rho} = B_1 u^2 + C_1 u + D_1, \tag{20}$$
$$\frac{\partial g}{\partial \rho} = B_3 v^2 + C_3 v + D_3.$$

From these expressions we calculate H by (5), obtaining

$$H = -(C_1 + C_3) + 2 \cdot \frac{C_1 + C_3 + (B_1 - D_3)u + (B_3 - D_1)v}{1 + uv}. \tag{21}$$

It is seen at once that in order that either $\partial H/\partial u$ or $\partial H/\partial v$ shall vanish identically, H must also be identically zero; a case evidently to be excluded, since then (1) defines a single surface only, and not a family. Therefore substituting (21) in (4), we obtain

$$\xi = \frac{(1 + uv)^2}{\dfrac{\partial H}{\partial u}} \int \frac{\partial}{\partial v}\left(\frac{\eta}{(1 + uv)^2} \cdot \frac{\partial H}{\partial v}\right) du,$$

and giving v a constant value for which η is holomorphic and $\partial H/\partial u$ does not vanish, it is evident that this expression for ξ has only a finite number of singularities, and a similar conclusion applies to η. Our proposition is thus proved in all three cases.

We now proceed to deduce certain functional equations connecting $\partial f/\partial \rho$ and ξ with $\partial g/\partial \rho$ and η, and thereby obtain a second classification of possible solutions.

Equation (6) contains only the four functions referred to (and their derivatives), and since for a given value of ρ either $\partial f/\partial \rho$ and ξ have only isolated singular points, or else $\partial g/\partial v$ and η have the same property, it is evidently possible to find two points u_0 and v_0 for which $1 + u_0 v_0 = 0$ and such that $\partial f/\partial \rho$ and ξ are holomorphic for $|u - u_0| < 2r$, $\partial g/\partial \rho$ and η holomorphic for $|v - v_0| < 2r$, where r is sufficiently small.*

Consequently, for $|u - u_0| < r$ and $|h|$ sufficiently small, the value of v defined by

$$v = -\frac{1}{u} + h,$$

will be such that $|v - v_0| < 2r$, and we may substitute this value of v in (6) and expand both sides in powers of h. This gives

* Since one of the two points u_0, v_0 is inside and the other outside the unit circle, this statement would be erroneous if, for instance, both f and g existed inside the unit circle only. Hence the necessity of the preceding investigation.

$$\cdots + 8h\xi(u, \rho)\frac{\partial^2 f(u,\rho)}{\partial u \partial \rho} + \frac{2h}{u}\frac{\partial f(u,\rho)}{\partial \rho}\frac{\partial \xi(u,\rho)}{\partial u} + 8\left(\frac{1}{u^3} - \frac{3h}{u^2}\right)\xi(u,\rho)\frac{\partial f(u,\rho)}{\partial \rho}$$

$$- 2uh\frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho}\frac{\partial \xi(u,\rho)}{\partial u} + 8\left(-\frac{1}{u} + h\right)\xi(u,\rho)$$

$$\times \left[\frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho} + h\frac{\partial^2 g\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)\partial \rho} + \cdots\right]$$

$$= \cdots + 8u^3 h\eta\left(-\frac{1}{u}, \rho\right)\frac{\partial^2 g\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)\partial \rho} + 2u^3 h\frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho}\frac{\partial \eta\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)}$$

$$- 8u^3\left[\eta\left(-\frac{1}{u}, \rho\right) + \frac{\partial \eta\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)}h + \cdots\right]$$

$$\times \left[\frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho} + \frac{\partial^2 g\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)\partial \rho}h + \cdots\right]$$

$$- 2uh\frac{\partial f(u,\rho)}{\partial \rho}\frac{\partial \eta\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)} + 8u\frac{\partial f(u,\rho)}{\partial \rho}$$

$$\times \left[\eta\left(-\frac{1}{u}, \rho\right) + \frac{\partial \eta\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)}h + \cdots\right],$$

where the terms of second and higher order in h are not written out.

Identifying the terms independent of h on both sides, we obtain

$$\left[\frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho} - \frac{1}{u^2}\frac{\partial f(u,\rho)}{\partial \rho}\right]\left[\eta\left(-\frac{1}{u}, \rho\right) - \frac{1}{u^4}\xi(u,\rho)\right] = 0. \tag{22}$$

Identifying the coefficients of h on both sides, and replacing differentiations in respect to $-(1/u)$ by such in respect to u, so that

$$\frac{\partial\eta\left(-\frac{1}{u},\rho\right)}{\partial\left(-\frac{1}{u}\right)}=u^2\frac{\partial\eta\left(-\frac{1}{u},\rho\right)}{\partial u},\qquad\frac{\partial^2 g\left(-\frac{1}{u},\rho\right)}{\partial\left(-\frac{1}{u}\right)\partial\rho}=u^2\frac{\partial^2 g\left(-\frac{1}{u},\rho\right)}{\partial u\partial\rho},$$

we find

$$\begin{aligned}&\frac{4}{u}\xi(u,\rho)\frac{\partial^2 f(u,\rho)}{\partial u\partial\rho}+\frac{1}{u}\frac{\partial f(u,\rho)}{\partial\rho}\frac{\partial\xi(u,\rho)}{\partial u}-\frac{12}{u^2}\xi(u,\rho)\frac{\partial f(u,\rho)}{\partial\rho}\\&\qquad-u\frac{\partial g\left(-\frac{1}{u},\rho\right)}{\partial\rho}\frac{\partial\xi(u,\rho)}{\partial u}-4u\xi(u,\rho)\frac{\partial^2 g\left(-\frac{1}{u},\rho\right)}{\partial u\partial\rho}\\&\qquad+4\xi(u,\rho)\frac{\partial g\left(-\frac{1}{u},\rho\right)}{\partial\rho}+3u^5\frac{\partial g\left(-\frac{1}{u},\rho\right)}{\partial\rho}\frac{\partial\eta\left(-\frac{1}{u},\rho\right)}{\partial u}\\&\qquad-3u^3\frac{\partial f(u,\rho)}{\partial\rho}\frac{\partial\eta\left(-\frac{1}{u},\rho\right)}{\partial u}=0.\end{aligned}\tag{23}$$

Supposing the first factor in (22) equal to zero, we have

$$\frac{\partial^2 g\left(-\frac{1}{u},\rho\right)}{\partial u\partial\rho}=\frac{1}{u^2}\frac{\partial^2 f(u,\rho)}{\partial u\partial\rho}-\frac{2}{u^3}\frac{\partial f(u,\rho)}{\partial\rho}$$

and (23) reduces to an identity.

On the other hand, when the second factor in (22) vanishes, but not the first, we have

$$\frac{\partial\eta\left(-\frac{1}{u},\rho\right)}{\partial u}=\frac{1}{u^4}\frac{\partial\xi(u,\rho)}{\partial u}-\frac{4}{u^5}\xi(u,\rho)$$

and (23) becomes

$$\begin{aligned}&\frac{\partial\xi(u,\rho)}{\partial u}\left[u^2\frac{\partial g\left(-\frac{1}{u},\rho\right)}{\partial\rho}-\frac{\partial f(u,\rho)}{\partial\rho}\right]\\&\qquad-2\xi(u,\rho)\left[u^2\frac{\partial^2 g\left(-\frac{1}{u},\rho\right)}{\partial u\partial\rho}+2u\frac{\partial g\left(-\frac{1}{u},\rho\right)}{\partial\rho}-\frac{\partial^2 f(u,\rho)}{\partial u\partial\rho}\right]=0.\end{aligned}$$

Since the first factor in (22) does not vanish, this may be written

$$\frac{1}{2}\frac{\partial\log\xi(u,\rho)}{\partial u}=\frac{\partial}{\partial u}\log\left[u^2\frac{\partial g\left(-\frac{1}{u},\rho\right)}{\partial\rho}-\frac{\partial f(u,\rho)}{\partial\rho}\right],$$

whence by integration

$$(24) \qquad u^2 \frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho} - \frac{\partial f(u, \rho)}{\partial \rho} = \Omega(\rho) \cdot \sqrt{\xi(u, \rho)},$$

where $\Omega(\rho) \neq 0$.

According to (22) and (24), we may now distinguish two cases: A and B.

Case A:

$$(25) \qquad \frac{\partial f(u, \rho)}{\partial \rho} - u^2 \frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho} = 0.$$

Substituting v for $-(1/u)$, this may obviously also be written

$$(25)' \qquad \frac{\partial g(v, \rho)}{\partial \rho} - v^2 \frac{\partial f\left(-\frac{1}{v}, \rho\right)}{\partial \rho} = 0.$$

From these equations an important conclusion may be drawn at once. Let us start at the point u_0 with a certain branch of the function $f(u, \rho)$ and continue it analytically along a closed path, returning to u_0 with a branch $f_1(u, \rho)$ of our function; then, since (6) subsists also for the branch f_1, and the continuation process does not affect the function g, it follows that (25) subsists also for the branch f_1, so that

$$\frac{\partial f_1(u, \rho)}{\partial \rho} = \frac{\partial f(u, \rho)}{\partial \rho},$$

and a similar argument applies to (25)′. *Therefore $\partial f(u, \rho)/\partial \rho$ and $\partial g(v, \rho)/\partial \rho$ are uniform functions of u and v.* The case A will be examined in § 4; it will be found convenient to subdivide it into three cases AI, AII and $AIII$ corresponding to the three cases of our first classification.

Case B:

$$(26) \qquad \eta\left(-\frac{1}{u}, \rho\right) - \frac{1}{u^4}\xi(u, \rho) = 0,$$

$$(24) \qquad u^2 \frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho} - \frac{\partial f(u, \rho)}{\partial \rho} = \Omega(\rho) \sqrt{\xi(u, \rho)}, \quad \Omega(\rho) \neq 0.$$

The last equation determines, rather completely, the form of $\xi(u, \rho)$. In fact, differentiating it in respect to u, we have

$$2u \frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho} + u^2 \frac{\partial^2 g\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)\partial \rho} \cdot \frac{1}{u^2} - \frac{\partial^2 f(u, \rho)}{\partial u \partial \rho} = \Omega(\rho) \frac{\partial \sqrt{\xi}}{\partial u},$$

or replacing $\frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho}$ by its value from (24),

$$\frac{\partial^2 g\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)\partial \rho} = \frac{\partial^2 f(u, \rho)}{\partial u \partial \rho} - \frac{2}{u}\frac{\partial f(u, \rho)}{\partial \rho} + \Omega(\rho)\left[\frac{\partial \sqrt{\xi}}{\partial u} - \frac{2}{u}\sqrt{\xi}\right].$$

Differentiating again and multiplying by u^2, we find

$$\frac{\partial^3 g\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)^2\partial \rho} = u^2 \frac{\partial^3 f(u, \rho)}{\partial u^2 \partial \rho} - 2u \frac{\partial^2 f(u, \rho)}{\partial u \partial \rho} + 2\frac{\partial f(u, \rho)}{\partial \rho}$$

$$+ \Omega(\rho)\left[u^2 \frac{\partial^2 \sqrt{\xi}}{\partial u^2} - 2u \frac{\partial \sqrt{\xi}}{\partial u} + 2\sqrt{\xi}\right],$$

and repeating the same procedure

$$\frac{\partial^4 g\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)^3\partial \rho} = u^4 \frac{\partial^4 f(u, \rho)}{\partial u^3 \partial \rho} + \Omega(\rho) u^4 \frac{\partial^3 \sqrt{\xi}}{\partial u^3};$$

but from (8) and (26) we obtain

$$\frac{\partial^3 g\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)^3} = u^4 \frac{\partial^3 f(u, \rho)}{\partial u^3};$$

differentiating in respect to ρ and comparing with the preceding equation, it is finally seen that

$$\frac{\partial^3 \sqrt{\xi}}{\partial u^3} = 0 \qquad \text{or} \qquad \sqrt{\xi} = au^2 + bu + c,$$

where a, b and c may depend on ρ. We can now replace (26) by the following

$$\xi(u, \rho) = (au^2 + bu + c)^2,$$

(27)

$$\eta(v, \rho) = v^4 \xi\left(-\frac{1}{v}, \rho\right) = (cv^2 - bv + a)^2.$$

The process of analytic continuation used in case A does not here show the uniformity of $\partial f/\partial \rho$, since in (24) $\Omega(\rho)$ may have different values for different branches of f. We may however conclude that, when $f(u, \rho)$ and $f_1(u, \rho)$ are two branches of f, there subsists a relation of the form

$$(28)\qquad \frac{\partial f_1(u, \rho)}{\partial \rho} - \frac{\partial f(u, \rho)}{\partial \rho} = h(\rho)\sqrt{\xi(u, \rho)} = h(\rho)(au^2 + bu + c).$$

The condition (26), which may be written

$$(26)'\qquad f'''(u) = \frac{1}{u^4} g'''\left(-\frac{1}{u}\right), \qquad g'''(v) = \frac{1}{v^4} f'''\left(-\frac{1}{v}\right),$$

has a simple geometric interpretation. From (1) we obtain

$$dx + idy = -u^2 f'''(u)du + g'''(v)dv,$$
$$dx - idy = f'''(u)du - v^2 g'''(v)dv,$$
$$dz = uf'''(u)du + vg'''(v)dv,$$

and writing the same equations for the parameter values u_1 and v_1,

$$dx_1 + idy_1 = -u_1^2 f'''(u_1)du_1 + g'''(v_1)dv_1,$$
$$dx_1 - idy_1 = f'''(u_1)du_1 - v_1^2 g'''(v_1)dv_1,$$
$$dz_1 = u_1 f'''(u_1)du_1 + v_1 g'''(v_1)dv_1.$$

Now make $u_1 = -1/v$, $v_1 = -1/u$, which gives

$$dx_1 + idy_1 = -\frac{1}{v^4} f'''\left(-\frac{1}{v}\right)dv + \frac{1}{u^2} g'''\left(-\frac{1}{u}\right)du$$
$$= -g'''(v)dv + u^2 f'''(u)du$$

by (26)′, so that $dx_1 + idy_1 = -(dx + idy)$. Similarly $(dx_1 - idy_1) = -(dx - idy)$, $dz_1 = -dz$, whence $dx_1 = -dx$, $dy_1 = -dy$, $dz_1 = -dz$, and denoting the constants of integration by $2x_0$, $2y_0$, $2z_0$, we have

$$x_1 - x_0 = -(x - x_0),\ y_1 - y_0 = -(y - y_0),\ z_1 - z_0 = -(z - z_0),$$

so that *the surface (1) is symmetrical about the point* x_0, y_0, z_0.

The case B will be examined in detail in § 5.

4. Case A. Detailed discussion.

Case AI: This case is quickly disposed of; for by (12) and (25)′ we have $\partial f/\partial \rho = A_0' + A_1'u + A_2'u^2$, $\partial g/\partial \rho = A_2' - A_1'v + A_0'v^2$, and the comparison with (20) and (21) gives $H = 0$, which was previously recognized as impossible.

Case AII: In eq. (13), let $u = u_1$ be a root of $P_4(u, \rho) = 0$; in the

vicinity of u_1 we then have an expansion

$$\frac{\partial^3 f}{\partial u^3} = \frac{a_\lambda}{(u-u_1)^\lambda} + \frac{a_{\lambda-1}}{(u-u_1)^{\lambda-1}} + \cdots + \frac{a_1}{u-u_1} + \text{positive powers,}$$
$$a_\lambda \neq 0 \text{ and } 1 \leqq \lambda \leqq 4.$$

If $u_1, a_1, \cdots, a_\lambda$ are independent of ρ, and if this is true for all the roots of $P_4(u, \rho) = 0$, then P_4 is independent of ρ and we find, as in case AI, $\partial f/\partial\rho = A_0 + A_1 u + A_2 u^2$, $\partial g/\partial\rho = A_2 - A_1 v + A_0 v^2$ and $H = 0$.

For the root u_1, we may therefore suppose that either u_1 itself depends on ρ, or in the contrary case, that at least one of $a_1, \cdots, a_\lambda$ is a function of ρ. From the above expansion we obtain

$$\frac{\partial^4 f}{\partial u^3 \partial\rho} = \frac{\lambda a_\lambda \dfrac{du_1}{d\rho}}{(u-u_1)^{\lambda+1}} + \frac{\dfrac{da_\lambda}{d\rho} + (\lambda-1)a_{\lambda-1}\dfrac{du_1}{d\rho}}{(u-u_1)^\lambda} + \cdots + \frac{\dfrac{da_2}{d\rho} + a_1\dfrac{du_1}{d\rho}}{(u-u_1)^2} \tag{29}$$
$$+ \frac{\dfrac{da_1}{d\rho}}{u-u_1} + \text{positive powers of } u - u_1.$$

On the other hand it is necessary, in order that $\partial f/\partial\rho$ be a uniform function of u, that the coefficients of $1/(u-u_1)$, $1/(u-u_1)^2$, $1/(u-u_1)^3$ in (29) all vanish. Consequently $\lambda = 3$ or 4; if $\lambda = 3$, then $du_1/d\rho \neq 0$, and we have another root u_2, which is simple and independent of ρ, as well as the coefficient of $1/(u-u_2)$ in the partial fraction for $\partial^3 f/\partial u^3$. We therefore have, in this case,

$$\xi(u, \rho) = \frac{(u-u_1)^3(u-u_2)}{k(\rho)}, \tag{30}$$

and decomposing its reciprocal into partial fractions

$$\frac{\partial^3 f(u, \rho)}{\partial u^3} = \frac{k(\rho)}{(u_1-u_2)^3}\left[\frac{(u_1-u_2)^2}{(u-u_1)^3} - \frac{u_1-u_2}{(u-u_1)^2} + \frac{1}{u-u_1} - \frac{1}{u-u_2}\right].$$

Since the coefficient of $1/(u-u_2)$ must be independent of ρ, we have $k(\rho) = -2c(u_1-u_2)^3$, where c is a constant, and we obtain

$$\frac{\partial f}{\partial\rho} = \frac{c(u_1-u_2)^2\dfrac{du_1}{d\rho}}{u-u_1} + A_0'(\rho) + A_1'(\rho)u + A_2'(\rho)u^2. \tag{31}$$

If however $\lambda = 4$, we may write

$$\xi(u, \rho) = \frac{(u-u_1)^4}{-6k(\rho)} \tag{32}$$

and obtain immediately

$$\text{(33)} \qquad \frac{\partial f}{\partial \rho} = \frac{k\dfrac{du_1}{d\rho}}{(u - u_1)^2} + \frac{\dfrac{dk}{d\rho}}{u - u_1} + A_0' + A_1'u + A_2'u^2.$$

Similar considerations apply to g, and it follows from (25) that $v_1 = -1/u_1$ and that $\partial f/\partial \rho$ and $\partial g/\partial \rho$ must both be of the form (31) or both of the form (33). In the former case,

$$\eta(v, \rho) = \frac{(v - v_1)^3(v - v_2)}{-2c_1(v_1 - v_2)^3},$$

$$\frac{\partial g}{\partial \rho} = \frac{c_1(v_1 - v_2)^2\dfrac{dv_1}{d\rho}}{v - v_1} + B_0' + B_1'v + B_2'v^2,$$

and computing H, we find

$$H = -\frac{2c(u_1 - u_2)^2\dfrac{du_1}{d\rho}}{(u - u_1)^2} - \frac{2c_1(v_1 - v_2)^2\dfrac{dv_1}{d\rho}}{(v - v_1)^2} - \frac{2v}{1 + uv}\cdot\frac{c(u_1 - u_2)^2\dfrac{du_1}{d\rho}}{u - u_1}$$

$$- \frac{2u}{1 + uv}\cdot\frac{c_1(v_1 - v_2)^2\dfrac{dv_1}{d\rho}}{v - v_1} - A_1' - B_1'$$

$$+ 2\frac{A_1' + B_1' + (A_2' - B_0')u + (B_2' - A_0')v}{1 + uv}.$$

In the expression

$$\frac{\partial}{\partial u}\left(\frac{\xi}{(1 + uv)^2}\frac{\partial H}{\partial u}\right)$$

there is one term containing $1/(v - v_1)$ as a factor, namely

$$\frac{c_1(v_1 - v_2)^2\dfrac{dv_1}{d\rho}}{c(u_1 - u_2)^3}\cdot\frac{\partial}{\partial u}\left(\frac{(u - u_1)^3(u - u_2)}{(1 + uv)^4}\right)\cdot\frac{1}{v - v_1},$$

whereas the right-hand side of (4), substituting H for θ, contains no negative powers of $v - v_1$, owing to the presence of the factor $(v - v_1)^3$ in $\eta(v, \rho)$. Consequently we must have

$$\frac{\partial}{\partial u}\left(\frac{(u - u_1)^3(u - u_2)}{(1 + uv)^4}\right) = 0 \text{ for } v = v_1 = -\frac{1}{u_1},$$

which condition immediately reduces to $u_1^4(u_1 - u_2) = 0$, but neither of these two factors can vanish on account of $du_1/d\rho \neq 0$, so that the case

considered is impossible. We now take up the case $\lambda = 4$; then ξ and $\partial f/\partial\rho$ are given by (32) and (33), and we also have

$$
\begin{aligned}
\eta(v,\,\rho) &= \frac{(v - v_1)^4}{-\,6k_1(\rho)}, \\
\frac{\partial g}{\partial\rho} &= \frac{k_1\dfrac{dv_1}{d\rho}}{(v - v_1)^2} + \frac{\dfrac{dk_1}{d\rho}}{v - v_1} + B_0' + B_1'v + B_2'v^2.
\end{aligned} \tag{34}
$$

Suppose first that $du_1/d\rho \neq 0$; then also $dv_1/d\rho \neq 0$ since $v_1 = -\,1/u_1$, and we find

$$
\begin{aligned}
H = {} & \frac{-\,2k\dfrac{du_1}{d\rho}}{(u - u_1)^3} - \frac{\dfrac{dk}{d\rho}}{(u - u_1)^2} - \frac{2k_1\dfrac{dv_1}{d\rho}}{(v - v_1)^2} - \frac{\dfrac{dk_1}{d\rho}}{v - v_1} \\
& - \frac{2v}{1 + uv}\left[\frac{k\dfrac{du_1}{d\rho}}{(u - u_1)^2} + \frac{\dfrac{dk}{d\rho}}{u - u_1}\right] - \frac{2u}{1 + uv}\left[\frac{k_1\dfrac{dv_1}{d\rho}}{(v - v_1)^2} + \frac{\dfrac{dk_1}{d\rho}}{v - v_1}\right] \\
& - A_1' - B_1' + 2\,\frac{A_1' + B_1' + (A_2' - B_0')u + (B_2' - A_0')v}{1 + uv}.
\end{aligned} \tag{35}
$$

Substituting this expression in (4), we see as before that, $\eta(v,\,\rho)$ containing the factor $(v - v_1)^4$, no negative powers of $v - v_1$ occur on the right-hand side of the equation, while on the left-hand side we have the terms

$$
\begin{aligned}
& \frac{1}{3k}\frac{\partial}{\partial u}\left(\frac{(u - u_1)^4}{(1 + uv)^4}\right)\cdot\left[\frac{k_1\dfrac{dv_1}{d\rho}}{(v - v_1)^2} + \frac{\dfrac{dk_1}{d\rho}}{v - v_1}\right] \\
& \qquad = \frac{4}{3k}\frac{(u - u_1)^3}{(1 + uv)^5}\left[\frac{k_1\dfrac{du_1}{d\rho}}{1 + u_1v} + u_1\frac{dk_1}{d\rho}\right],
\end{aligned} \tag{36}
$$

where, in the last expression, we have replaced v_1 by $-\,1/u_1$. It is evident that the factor multiplying $1/(1 + u_1v)$ does not vanish for $v = -\,1/u_1$, so that also this case is impossible.

Next, we assume that $du_1/d\rho = 0$ but $u_1 \neq 0$; then we must have $dk/d\rho \neq 0$. The expressions for $\partial f/\partial\rho$ and $\partial g/\partial\rho$ now become, observing that $v_1 = -\,1/u_1$ and writing k' and k_1' for $dk/d\rho$ and $dk_1/d\rho$,

$$
\begin{aligned}
\frac{\partial f}{\partial\rho} &= \frac{k'}{u - u_1} + A_0' + A_1'u + A_2'u^2, \\
\frac{\partial g}{\partial\rho} &= \frac{u_1k_1'}{1 + u_1v} + B_0' + B_1'v + B_2'v^2.
\end{aligned} \tag{37}
$$

Substituting these expressions in (25), we obtain

$$
(38) \qquad \begin{aligned}
k' &= u_1^4 k_1', \\
k_1' u_1^3 + B_2' - A_0' &= 0, \\
k_1' u_1^2 - B_1' - A_1' &= 0, \\
k_1' u_1 + B_0' - A_2' &= 0,
\end{aligned}
$$

and the expression for H becomes

$$
(39) \qquad \begin{aligned}
\frac{1}{u_1 k_1'} \cdot H = &- \frac{u_1^3}{(u - u_1)^2} - \frac{u_1}{(1 + u_1 v)^2} - \frac{2v}{1 + uv} \cdot \frac{u_1^3}{u - u_1} \\
&- \frac{2u}{1 + uv} \cdot \frac{1}{1 + u_1 v} - u_1 + 2 \frac{u_1 + u - u_1^2 v}{1 + uv}.
\end{aligned}
$$

Introducing this in (4), we obtain after some algebraic reductions

$$
- \frac{1}{k} \cdot \frac{u_1^3}{(1 + uv)^2} = - \frac{1}{k_1 u_1^4} \cdot \frac{u_1^3}{(1 + uv)^2},
$$

whence $k = k_1 u_1^4$. Writing k instead of k_1, and expressing the B' in terms of the A' by (38), equations (37) give upon integration in respect to ρ

$$
(40) \qquad \begin{aligned}
f(u, \rho) &= \frac{u_1^4 k(\rho)}{u - u_1} + A_0(\rho) + A_1(\rho) u + A_2(\rho) u^2, \\
g(v, \rho) &= \frac{u_1 k(\rho)}{1 + u_1 v} + A_2(\rho) - u_1 k(\rho) + [u_1^2 k(\rho) - A_1(\rho)] v \\
&\qquad + [A_0(\rho) - u_1^3 k(\rho)] v^2.
\end{aligned}
$$

The family of minimal surfaces defined by these values of f and g is a special case of one determined later (case $B2$).

Finally, we have to discuss the case when $u_1 = 0$; then

$$
\frac{\partial^3 f}{\partial u^3} = \frac{- 6k(\rho)}{u^4}; \qquad f = \frac{k}{u} + A_0 + A_1 u + A_2 u^2,
$$

$$
\frac{\partial f}{\partial \rho} = \frac{k'}{u} + A_0' + A_1' u + A_2' u^2;
$$

$$
\frac{\partial g}{\partial \rho} = v^2 \frac{\partial f\left(- \frac{1}{v}, \rho\right)}{\partial \rho} = - k' v^3 + A_2' - A_1' v + A_0' v^2,
$$

$$
g(v, \rho) = - k v^3 + A_2 - A_1 v + A_0 v^2 + g_0(v),
$$

$$
\frac{\partial^3 g}{\partial v^3} = - 6k(\rho) + \frac{d^3 g_0(v)}{dv^3},
$$

$$\frac{1}{k'}H = -\frac{1}{u^2} - 3v^2 - \frac{2v}{1+uv}\cdot\frac{1}{u} + \frac{2uv^3}{1+uv}$$

$$= -\frac{1}{u^2} - v^2 - 2\frac{v}{u},$$

and substituting this value of H in (4), we find that this equation reduces to

$$\frac{d^3g_0(v)}{dv^3} = 0,$$

so that we obtain the following expressions for f and g,

$$(41)\qquad \begin{aligned} f(u,\,\rho) &= \frac{k(\rho)}{u} + A_0(\rho) + A_1(\rho)u + A_2(\rho)u^2, \\ g(v,\,\rho) &= -\,k(\rho)v^3 + A_2(\rho) - A_1(\rho)v + A_0(\rho)v^2. \end{aligned}$$

We have here assumed $g_0(v) = 0$, which is legitimate, since the addition to $g(v,\rho)$ of a polynomial of the second degree with coefficients independent of ρ merely signifies a simultaneous translation, of constant magnitude and direction, of the entire family (1). The family defined by (41) is again a special case of one considered later (case *B3*).

Case AIII: We begin by examining the cases where, in (16) and (18),

$$(42)\qquad ad - bc = 0.$$

First, suppose $a = b = c = d = 0$; then (20) and (25)′ give $\partial f/\partial\rho = B_1u^2 + C_1u + D_1$, $\partial g/\partial\rho = D_1v^2 - C_1v + B_1$, and by (21) $H = 0$, which was seen to be impossible. Second, suppose $b = d = 0$, but either $a \neq 0$ or $c \neq 0$; then the second equation (18) gives

$$\frac{\partial g}{\partial\rho} = A_4v^2 + B_4v + C_4 + \frac{D_4}{v},$$

whence, by (25),

$$\frac{\partial f}{\partial\rho} = A_4 - B_4u + C_4u^2 - D_4u^3,$$

and according as $a \neq 0$ or $a = 0$, the first or second equation (16) gives $\xi = P_4(u,\rho)$ or $\xi = P_5(u,\rho)$. But at a zero of ξ, $\partial f/\partial\rho$ becomes infinite according to (8) and thus contradicts the first equation (16), unless this zero and all the corresponding coefficients in the partial fraction for $1/\xi$ are independent of ξ, that is, unless $\partial\xi/\partial\rho = 0$. Then, however, (8) and (25)′ give $\partial f/\partial\rho = A_0' + A_1'u + A_2'u^2$, $\partial g/\partial\rho = A_2' - A_1'v + A_0'v^2$, and we have again $H = 0$. The case $c = d = 0$ is treated in the same way. The case $d = 0$, $b \neq 0$ and $c \neq 0$ is impossible, since then $ad - bc \neq 0$.

There now remains the case $d \neq 0$, in which we obtain from (16), (18) and (42)

$$(43) \qquad (bu - d)\frac{\partial f}{\partial \rho} = (bA_2 - dA_1)u^3 + \cdots + (bD_2 - dD_1),$$

$$(44) \qquad (cv - d)\frac{\partial g}{\partial \rho} = (cA_4 - dA_3)v^3 + \cdots + (cD_4 - dD_3),$$

and applying (25) to (44), we find

$$(45) \qquad (du + c)\frac{\partial f}{\partial \rho} = (cA_4 - dA_3)\cdot\frac{1}{u} - (cB_4 - dB_3) + (cC_4 - dC_3)u - (cD_4 - dD_3)u^2.$$

If $b = 0$, then (42) gives $a = 0$, (43) and (45) show that $A_1 = 0$, so that the first of (16) gives $\partial f/\partial \rho$ as a polynomial of the second degree in u, whence we conclude as before that $H = 0$. If $b \neq 0$, and $bu - d$ is a factor of the right-hand side in (43), we arrive at the same conclusion. Finally, when the right-hand side in (43) is not divisible by $bu - d$, the comparison of (43) and (45) gives

$$b/d = -d/c \qquad \text{or} \qquad bc + d^2 = 0,$$

$$cA_4 - dA_3 = 0, \qquad bA_2 - dA_1 = 0,$$

whence, by (43),

$$(46) \qquad \frac{\partial f}{\partial \rho} = \frac{P_2(u, \rho)}{bu - d}.$$

The combination of $bc + d^2 = 0$ with (42) gives $a + d = 0$, and by (46), the first of equations (16) may now be written

$$(bu - d)\frac{\partial \xi}{\partial u} - 4b\xi = \frac{P_2(u, \rho)}{bu - d} - A_1u^3 - B_1u^2 - C_1u - D_1.$$

Integrating this linear differential equation of the first order, we obtain

$$\xi = (bu - d)^4\left[\int\left(\frac{P_2(u, \rho)}{(bu - d)^5} - \frac{A_1u^3 + B_1u^2 + C_1u + D_1}{(bu - d)^4}\right)du + C(\rho)\right],$$

or performing the integration

$$\xi = P_4(u, \rho) - \frac{A_1}{b^3}(bu - d)^4 \log (bu - d).$$

Substituting this expression back into the first equation (16), we see that, since $a = -d$, $\partial f/\partial \rho = P_3(u, \rho)$, so that $\partial f/\partial \rho$ has no pole at $-d/b$ contrary to our hypothesis.

We may now assume that $ad - bc \neq 0$. Then equations (16) may

be solved for ξ and $\partial\xi/\partial u$, giving

$$
(47)\quad
\begin{aligned}
\xi(u, \rho) &= -\frac{1}{4(ad-bc)}\Big\{[bu^2+(a-d)u-c]\frac{\partial f}{\partial \rho} - (a+bu) \\
&\qquad \times (A_2u^3+\cdots+D_2) + (c+du)(A_1u^3+\cdots+D_1)\Big\}, \\
\frac{\partial \xi(u, \rho)}{\partial u} &= \frac{1}{ad-bc}\Big\{(d-bu)\frac{\partial f}{\partial \rho} - d(A_1u^3+\cdots+D_1) \\
&\qquad + b(A_2u^3+\cdots+D_2)\Big\},
\end{aligned}
$$

whence, differentiating the first equation in respect to u and substituting in the second, we obtain the following linear differential equation for $\partial f/\partial\rho$:

$$
(48)\quad [bu^2+(a-d)u-c]\frac{\partial^2 f}{\partial u\partial\rho} + [-2bu+(a+3d)]\frac{\partial f}{\partial\rho} + P_2(u, \rho) = 0,
$$

where the coefficients in the polynomial of the second degree in u, $P_2(u, \rho)$, are of no particular interest to us. From (18), we get expressions for η and $\partial\eta/\partial v$ similar to (47), and the differential equation for $\partial g/\partial\rho$:

$$
(49)\quad [cv^2+(a-d)v-b]\frac{\partial^2 g}{\partial v\partial\rho} + [-2cv+(a+3d)]\frac{\partial g}{\partial v} + Q_2(v, \rho) = 0.
$$

From (47), and the similar expression for $\eta(v, \rho)$ it follows that, since $\partial f(u, \rho)/\partial\rho$ and $\partial g(v, \rho)/\partial\rho$ are uniform functions of u and v respectively, the same is also the case with $\xi(u, \rho)$ and $\eta(v, \rho)$.

Let us first suppose that the roots α and β of $bu^2+(a-d)u-c=0$ are distinct; the roots $-1/\alpha$ and $-1/\beta$ of $cv^2+(a-d)v-b=0$ are then also distinct. According to (17), the only singularities of ξ are α and β (and infinity, if both α and β are finite). Suppose that α is finite, then, since α and β are distinct, α is a regular singular point for the linear differential equation (17), and since ξ is uniform, α can only be a pole, in the vicinity of which we have the expansion

$$
\xi = \frac{a_\lambda}{(u-\alpha)^\lambda} + \frac{a_{\lambda-1}}{(u-\alpha)^{\lambda-1}} + \cdots + \frac{a_1}{u-\alpha} + \cdots, \quad \lambda > 0,\ a_\lambda \neq 0.
$$

But at a pole of ξ, its reciprocal $\partial^3 f/\partial u^3$ is holomorphic, and consequently also $\partial f/\partial\rho$. Substituting the expansion of ξ in the first equation (16), we obtain

$$
a + b\alpha = 0, \qquad -\lambda b - 4b = 0,
$$

whence, since $\lambda > 0$, $b = a = 0$ and $ad - bc = 0$ contrary to our assumption. Thus ξ has no poles at finite distance, but is an entire function, and by the first of (16) the same is true of $\partial f/\partial\rho$. A similar reasoning shows

that $\partial g/\partial \rho$ is an entire function of v, and (25) and (25)$'$ prove that either function has a pole of no higher than the second order at infinity, that is, $\partial f/\partial \rho$ and $\partial g/\partial \rho$ are polynomials of the second degree. Now (25) and (21) finally give $H = 0$, which is impossible. As is readily seen, the preceding argument holds even when $bu^2 + (a - d)u - c$ vanishes identically.

There remains the case when $bu^2 + (a - d)u - c = 0$ has a finite double root α, so that

$$(a - d)^2 + 4bc = 0. \tag{50}$$

If ξ has a pole at α, the preceding argument holds; we must therefore suppose that α is an essential singularity of ξ. Then α is also an essential singularity of $\partial f/\partial \rho$, for in the contrary case, the first equation (16), regarded as a linear differential equation in ξ, would show that α were at most a pole of ξ.

Now transform (49) by making $v = -1/u$; as (25) gives

$$u^2 \frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho} = \frac{\partial f(u, \rho)}{\partial \rho},$$

$$2u \frac{\partial g\left(-\frac{1}{u}, \rho\right)}{\partial \rho} + \frac{\partial^2 g\left(-\frac{1}{u}, \rho\right)}{\partial\left(-\frac{1}{u}\right)\partial \rho} = \frac{\partial^2 f(u, \rho)}{\partial \rho},$$

we obtain

$$[bu^2 + (a - d)u - c]\frac{\partial^2 f}{\partial u \partial \rho} + [-2bu - (3a + d)]\frac{\partial f}{\partial \rho} - u^2 Q\left(-\frac{1}{u}, \rho\right) = 0.$$

This equation must be identical with (48), since otherwise the combination of both would show that $\partial f/\partial \rho$ were holomorphic at α; consequently

$$a + d = 0,$$

and this equation, combined with (50), gives $ad - bc = 0$ contrary to our hypothesis.

5. Case B. Detailed discussion.

This case will be subdivided according to the various forms which the expressions (27) for ξ and η can take, in the following manner

Case B1: $a \neq 0, \quad b^2 - 4ac \neq 0,$

Case B2 $a \neq 0, \quad b^2 - 4ac = 0, \quad c \neq 0,$

Case B3: $a \neq 0, \quad b^2 - ac = 0, \quad c = 0,$

Case B4: $a = c = 0.$

The only remaining possibility, $a = 0$ and $c \neq 0$, reduces, by interchanging u and f with v and g, to B1 or B3 according as $b \neq 0$ or $b = 0$.

Case B1.—We have, denoting by u_1 and u_2 the roots of $au^2 + bu + c = 0$,

$$\xi(u, \rho) = \frac{(u - u_1)^2(u - u_2)^2}{k}, \qquad u_1 \mp u_2,$$

and by (8) and (27)

$$(51) \qquad \begin{aligned} \frac{\partial^3 f}{\partial u^3} &= \frac{k}{(u - u_1)^2(u - u_2)^2}, \\ \frac{\partial^3 g}{\partial v^3} &= \frac{k}{(1 + u_1 v)^2(1 + u_2 v)^2}. \end{aligned}$$

Decomposing into partial fractions, we find

$$(52) \qquad \begin{aligned} \frac{\partial^3 f}{\partial u^3} &= \frac{k}{(u_1 - u_2)^2}\left[\frac{1}{(u - u_1)^2} + \frac{1}{(u - u_2)^2}\right] \\ &\qquad - \frac{2k}{(u_1 - u_2)^3}\left[\frac{1}{u - u_1} - \frac{1}{u - u^2}\right], \\ \frac{\partial^3 g}{\partial v^3} &= \frac{k}{(u_1 - u_2)^2}\left[\frac{u_1^2}{(1 + u_1 v)^2} + \frac{u_2^2}{(1 + u_2 v)^2}\right] \\ &\qquad - \frac{2ku_1u_2}{(u_1 - u_2)^3}\left[\frac{u_1}{1 + u_1 v} - \frac{u_2}{1 + u_2 v}\right], \end{aligned}$$

whence

$$\begin{aligned} \frac{\partial^4 f}{\partial u^3 \partial \rho} &= \frac{2k}{(u_1 - u_2)^2}\left[\frac{\frac{du_1}{d\rho}}{(u - u_1)^3} + \frac{\frac{du_2}{d\rho}}{(u - u_2)^3}\right] \\ &\quad + \left[\frac{\frac{dk}{d\rho}}{(u_1 - u_2)^2} - \frac{2k\left(2\frac{du_1}{d\rho} - \frac{du_2}{d\rho}\right)}{(u_1 - u_2)^3}\right] \cdot \frac{1}{(u - u_1)^2} \\ &\quad + \left[\frac{\frac{dk}{d\rho}}{(u_1 - u_2)^2} - \frac{2k\left(\frac{du_1}{d\rho} - 2\frac{du_2}{d\rho}\right)}{(u_1 - u_2)^3}\right] \cdot \frac{1}{(u - u_2)^2} \\ &\quad - 2\frac{d}{d\rho}\left(\frac{k}{(u_1 - u_2)^3}\right) \cdot \left[\frac{1}{u - u_1} - \frac{1}{u - u_2}\right] \end{aligned}$$

and

$$(53) \qquad \begin{aligned} \frac{\partial f}{\partial \rho} &= \frac{k}{(u_1 - u_2)^2}\left[\frac{du_1}{d\rho}\log(u - u_1) + \frac{du_2}{d\rho}\log(u - u_2)\right] \\ &\quad - \left[\frac{\frac{dk}{d\rho}}{(u_1 - u_2)^2} - \frac{2k\left(2\frac{du_1}{d\rho} - \frac{du_2}{d\rho}\right)}{(u_1 - u_2)^3}\right](u - u_1)\log(u - u_1) \end{aligned}$$

$$(53)\qquad - \left[\frac{\dfrac{dk}{d\rho}}{(u_1 - u_2)^2} - \frac{2k\left(\dfrac{du_1}{d\rho} - 2\dfrac{du_2}{d\rho}\right)}{(u_1 - u_2)^3} \right] (u - u_2) \log (u - u_2)$$

$$- \frac{d}{d\rho}\left(\frac{k}{(u_1 - u_2)^3}\right)[(u - u_1)^2 \log (u - u_1) - (u - u_2)^2 \log (u - u_2)]$$

$$+ A_0' + A_1'u + A_2'u^2.$$

Making an analytic continuation around $u = u_1$ and applying (28), we obtain

$$\frac{k}{(u_1 - u_2)^2}\frac{du_1}{d\rho} - [\cdots](u - u_1) - \frac{d}{d\rho}\left(\frac{k}{(u_1 - u_2)^3}\right)\cdot (u - u_1)^2$$
$$= h(u - u_1)(u - u_2),$$

and setting $u = u_1$, this reduces to $du_1/d\rho = 0$. Similarly $du_2/d\rho = 0$, so that u_1 and u_2 are independent of ρ. We may now show that this case reduces to case *B4* by changing our coördinate axes.

For this purpose we shall use a theorem which is obtained immediately by combining the formulas in Darboux's Théorie des Surfaces, vol. I, p. 391, with those p. 46 and following, namely:

The substitution of U, V, $F(U)$ and $G(V)$ for u, v, $f(u)$ and $g(v)$ in equations (1), where

$$(54)\qquad U = \frac{lu - m}{nu + p}, \qquad V = \frac{pv - n}{mv + l},$$
$$\frac{d^3F(U)}{dU^3} = \frac{(nu + p)^4}{(lp + mn)^2}\frac{d^3f(u)}{du^3}, \qquad \frac{d^3G(V)}{dV^3} = \frac{(mv + l)^4}{(lp + mn)^2}\frac{d^3g(v)}{dv^3},$$

and l, m, n, p are constants ($lp + mn \neq 0$), is equivalent to the orthogonal coördinate transformation defined by

$$X + iY = \frac{l^2(x + iy) - m^2(x - iy) + 2lmz}{lp + mn},$$

$$(55)\qquad X - iY = \frac{-n^2(x + iy) + p^2(x - iy) + 2npz}{lp + mn},$$

$$Z = \frac{-ln(x + iy) - mp(x - iy) + (lp - mn)z}{lp + mn}.$$

In the case under consideration, make $l = 1$, $m = u_1$, $n = 1$, $p = -u_2$; since u_1 and u_2 are independent of ρ, the same is the case with the coefficients in (55), so that every individual surface of our family is referred to the *same* new system of coördinates X, Y, Z. From (54) we obtain

$$U = \frac{u - u_1}{u - u_2}, \qquad V = -\frac{1 + u_2 v}{1 + u_1 v},$$

$$\frac{\partial^3 F}{\partial U^3} = \frac{(u - u_2)^4}{(u_1 - u_2)^2} \cdot \frac{k}{(u - u_1)^2 (u - u_2)^2} = \frac{k}{(u_1 - u_2)^2} \cdot \frac{1}{U^2},$$

$$\frac{\partial^3 G}{\partial V^3} = \frac{(1 + u_1 v)^4}{(u_1 - u_2)^2} \cdot \frac{k}{(1 + u_1 v)^2 (1 + u_2 v)^2} = \frac{k}{(u_1 - u_2)^2} \cdot \frac{1}{V^2},$$

or writing

$$K(\rho) = -\frac{2k(\rho)}{(u_1 - u_2)^2},$$

$$\frac{\partial^3 F}{\partial U^3} = -\frac{K(\rho)}{2U^2}, \qquad \frac{\partial^3 G}{\partial V^3} = -\frac{K(\rho)}{2V^2},$$

which are precisely the equations of definition of case *B4*.

Case B2.—We have

$$\frac{\partial^3 f(u, \rho)}{\partial u^3} = -\frac{6ku_1^4}{(u - u_1)^4}, \qquad \frac{\partial^3 g(v, \rho)}{\partial v^3} = -\frac{6ku_1^4}{(1 + u_1 v)^4}, \qquad u_1 \neq 0,$$

whence

$$f(u, \rho) = \frac{ku_1^4}{u - u_1} + A_0 + A_1 u + A_2 u^2,$$

$$g(v, \rho) = \frac{ku_1}{1 + u_1 v} + B_0 + B_1 v + B_2 v^2.$$

If first $du_1/d\rho \neq 0$, the reasoning which led to formulas (35) and (36) is still valid, and yields the same contradiction as in case *A*. Therefore u_1 is independent of ρ, and the present case may be reduced to case *B3* by making $l = 1$, $m = u_1$, $n = 0$, $p = 1$ in (54) and (55), whence

$$U = u - u_1, \qquad V = \frac{v}{1 + u_1 v},$$

$$\frac{\partial^3 F}{\partial U^3} = -\frac{6ku_1^4}{(u - u_1)^4} = -\frac{6ku_1^4}{U^4},$$

$$\frac{\partial^3 G}{\partial V^3} = -(1 + u_1 v)^4 \cdot \frac{6ku_1^4}{(1 + u_1 v)^4} = -6ku_1^4,$$

or by writing $K(\rho) = k(\rho) \cdot u_1^4$,

$$\frac{\partial^3 F}{\partial U^3} = -\frac{6K(\rho)}{U^4}, \qquad \frac{\partial^3 G}{\partial V^3} = -6K(\rho),$$

that is, the equations of definition of case *B3*.

Case B3.—We now have

$$\xi(u,\ \rho) = \frac{u^4}{-\ 6k(\rho)}, \qquad \eta(v,\ \rho) = -\frac{1}{6k(\rho)},$$

$$f(u,\ \rho) = \frac{k}{u} + A_0 + A_1 u + A_2 u^2,$$

$$g(v,\ \rho) = -\ kv^3 + B_0 + B_1 v + B_2 v^2,$$

and (24) gives, writing $\Omega(\rho)\sqrt{\xi(u,\ \rho)} = l'(\rho)\cdot u^2$,

$$B_0' - A_2' = l', \qquad B_1' + A_1' = 0, \qquad B_2' - A_0' = 0,$$

$$H = -\ k'\left(\frac{1}{u^2} + v^2 + 2\frac{v}{u}\right) - \frac{2l'u}{1 + uv}.$$

As is readily seen, this H satisfies (4); the corresponding family of minimal surfaces is

$$x + iy = -\ 6k(\rho)\left(\frac{1}{u} + v\right),$$

(56) $$x - iy = 2k(\rho)\left(\frac{1}{u^3} + v^3\right) - l(\rho),$$

$$z = 3k(\rho)\left(\frac{1}{u^2} - v^2\right).$$

None of these surfaces are real; in the special case $l = 0$, we obtain the surfaces defined by (41). The elimination of the parameters u and v in (56) gives

(57) $$\frac{1}{432k^2}(x + iy)^4 + (x + iy)(x - iy + l) + z^2 = 0.$$

The only singularity of this surface is the double point $x = -\ \frac{1}{2}l$, $y = -\ \frac{1}{2}li$, $z = 0$, which is also a point of symmetry of the surface. The only straight line on the surface is $x + iy = 0$, $z = 0$.

Case B4.—We may write

$$\xi(u,\ \rho) = -\frac{2u^2}{k(\rho)}, \qquad \eta(v,\ \rho) = -\frac{2v^2}{k(\rho)},$$

$$f(u,\ \rho) = \frac{k}{2}(u \log u - u) + A_0 + A_1 u + A_2 u^2,$$

$$g(v,\ \rho) = \frac{k}{2}(v \log v - v) + B_0 + B_1 v + B_2 v^2,$$

and (24) gives, $l(\rho)$ being arbitrary,

$$B_0' - A_2' = 0, \qquad B_1' + A_1' = -\ k' - l', \qquad B_2' - A_2' = 0.$$

The equation (4) is satisfied, and the corresponding minimal surfaces are catenoids with the z-axis as common axis of revolution; for (1) gives

$$
(58) \qquad
\begin{aligned}
x + iy &= \frac{k}{2}\left(u + \frac{1}{v}\right), \\
x - iy &= \frac{k}{2}\left(\frac{1}{u} + v\right), \\
z &= -\frac{k}{2}(\log u + \log v) + l,
\end{aligned}
$$

whence

$$
x^2 + y^2 = \frac{k^2}{4}\left(uv + 2 + \frac{1}{uv}\right), \qquad \sqrt{x^2 + y^2} = \frac{k}{2}\left(\sqrt{uv} + \frac{1}{\sqrt{uv}}\right),
$$

$$
\sqrt{uv} = e^{-(z-l)/k},
$$

$$
(59) \qquad \sqrt{x^2 + y^2} = \frac{k}{2}\left(e^{(z-l)/k} + e^{-(z-l)/k}\right),
$$

$k = k(\rho)$ and $l = l(\rho)$ being arbitrary functions of ρ.

PRINCETON, July 10, 1915.

AN ISOMORPHISM BETWEEN THETA CHARACTERISTICS AND THE $(2p+2)$-POINT.*

By Arthur B. Coble.†

The theta functions of two and three variables have been studied in connection with respectively the set of 6 points on a conic and the set of 8 base points of a net of quadrics. The main object of this article is to show that there is a grouping of the theta functions of p variables which is isomorphic with the grouping of the simplest system of irrational invariants of a set of $2p+2$ points in a projective space S_p. This isomorphism is set forth in § 2. Theta relations suggested by it are developed in § 4 for the case $p = 3$. In § 1 the tactical relations of the period and theta characteristics which appear most clearly in the light of finite geometry,‡ are developed in remarkably simple form with reference to a basis configuration. In § 3 a transition is made from the geometric to the arithmetic characteristic in order to determine the signs in a theta relation.

§ 1. The Characteristic Theory With Reference to a Basis Configuration.

In the finite space S_{2p-1} (mod. 2) let us take the supernumerary coördinates

$$(1) \qquad y_i \ (i = 1, \cdots, 2p+1), \qquad \textstyle\sum_{i=1}^{i=2p+1} y_i = 0.$$

Then the bilinear relation between the points y and y',

$$(2) \qquad y_1y_1' + y_2y_2' + \cdots + y_{2p+1}y'_{2p+1} = 0,$$

is the equation of a null system C. The points y and y' are *syzygetic* or *azygetic* according as they do or do not satisfy (2). The points of S_{2p-1} are identified with the period characteristics (other than the zero characteristic) of the theta functions, the null system C with the bilinear relation on the periods, and the projectivities in S_{2p-1} with the integral linear transformations on the periods when reduced modulo 2.

The $2p+1$ S_{2p-2}'s, $y_i = 0$, constitute an S_{2p-2}-*base* in the usual projective sense. As usual it determines a *point-base* ordered with regard

* Read before the American Mathematical Society, Jan. 1, 1915.

† This investigation is in progress under the auspices of the Carnegie Institution of Washington, D. C.

‡ Coble, "An application of finite geometry etc.," Transactions of the American Mathematical Society, vol. XIV (1913), p. 241; referred to hereafter as F. G.

to the S_{2p-2}-base. The point P_{0i} of the point-base associated with the S_{2p-2}, $y_i = 0$, has the coördinates, $y_i = 0$, $y_k = 1$ $(k \neq i)$. Evidently corresponding point and S_{2p-2} are point and null S_{2p-2} under C whence the two bases will be called a *self-dual base* of C and will be denoted by B_0. The further point, $P_{ik} = P_{ki} = P_{0i} + P_{0k}$ $(i \neq k; i, k = 1, \cdots, 2p + 1)$, on the line $P_{0i}P_{0k}$ has coördinates $y_i = y_k = 1$, $y_j = 0$ $(i \neq j \neq k)$ and is called a *residual* point of the base B_0. According to F. G., p. 271, the $(p + 1)(2p + 1)$ points P_{ik} $(i \neq k;\ i, k = 0, 1, \cdots, 2p + 1)$ form the basis points and the residual points not only of the particular base B_0 but also of any one of $2p + 2$ bases B_i $(i = 0, 1, \cdots, 2p + 1)$ where the base B_i contains the $2p + 1$ points P_{ik} for which $k \neq i$, $k = 0, 1, 2, \cdots, 2p + 1$. These bases B_i also are self dual under C, the null S_{2p-2} of P_{ik} being $y_i + y_k = 0$. Thus the $2p + 2$ bases B_i form a symmetrical *basis configuration* B and the point P_{ik} of B belongs to the bases B_i and B_k.

The remaining points of S_{2p-1} are determined in conformity with (1) by the following relations:

$$
\begin{aligned}
&P_{ij} + P_{jk} = P_{ik},\\
&\qquad P_{ijkl} = P_{ij} + P_{kl} = P_{ik} + P_{jl} = \cdots,\\
(3)\qquad &\qquad P_{ijklmn} = P_{ijkl} + P_{mn} = \cdots = P_{ij} + P_{kl} + P_{mn} = \cdots,\\
&\qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&P_{012\cdots(2k-1)} = P_{(2k)(2k+1)\cdots(2p+1)},\\
&P_{012\cdots(2p+1)} = 0.
\end{aligned}
$$

That the $2^{2p} - 1$ points of S_{2p-1} are thus given, each in two ways, by the above notation is verified by

$$\binom{2p+2}{2} + \binom{2p+2}{4} + \cdots + \binom{2p+2}{2p} = \frac{1}{2}\,2^{2p+2} - 2 = 2(2^{2p} - 1).$$

It is clear from the form of (2) that

THEOREM 1. *Two points of S_{2p-1} are syzygetic or azygetic according as a symbol of one has an even or an odd number of subscripts in common with a symbol of the other.*

It is shown in F. G., p. 271, that the group of C contains a subgroup $G_{(2p+2)!}$ which permutes symmetrically the bases B_i. The collineations which correspond to the transpositions in the symmetric group are the involutions which are attached to the points of B as described in F. G., p. 246 (12). Then one sees at once from the above notation for the points that

THEOREM 2. *Under the $G_{(2p+2)!}$ of B the points of S_{2p-1} divide into $p/2$*

or $(p+1)/2$ conjugate sets according as p is even or odd. The k-th set $(k = 1, 2, \cdots, p/2$ or $(p+1)/2)$ consists of the $\binom{2p+2}{2k}$ points whose symbol has $2k$ or $2p+2-2k$ subscripts; except in the case $k = (p+1)/2$ when the number is $\frac{1}{2}\binom{2p+2}{p+1}$.

Let us now classify the 2^{2p} quadrics which belong to C (*i. e.*, whose polar systems coincide with C) according to their behavior under $G_{(2p+2)!}$. According to F. G., p. 270 and p. 257, there is a unique quadric, $R_y = \sum y_i y_k$ $(i, k = 1, \cdots, 2p+1;\ i < k)$, which contains none of the residual points of B_0. The remaining quadrics are obtained from R_y by attaching any number of the squared terms $y_1^2, \cdots, y_{2p+1}^2$ to R_y. Since $\sum y_i^2 = 0$, any quadric can be expressed in two ways whence $\frac{1}{2}\, 2^{2p+1} = 2^{2p}$ quadrics are so obtained.

Consider the particular quadric, $R_y + y_1^2 + \cdots + y_p^2$. The point P_{ik} when either $i, k = 1, \cdots, p$ $(i < k)$ or $i, k = p+1, \cdots, 2p+1$ $(i < k)$ is not on R_y but is on $y_1^2 + \cdots + y_p^2$ and therefore is not on the quadric; while the point P_{ik} when $i = 1, \cdots, p$ and $k = p+1, \cdots, 2p+1$ is not on R_y nor on $y_1^2 + \cdots + y_p^2$ and therefore is on the quadric. The point P_{0k} is $\left\{\begin{matrix}\text{on}\\ \text{not on}\end{matrix}\right\}$ R_y if p is $\left\{\begin{matrix}\text{even}\\ \text{odd}\end{matrix}\right\}$; the point P_{0k} is $\left\{\begin{matrix}\text{on}\\ \text{not on}\end{matrix}\right\}$ $y_1^2 + \cdots + y_p^2$ either if p is $\left\{\begin{matrix}\text{odd}\\ \text{even}\end{matrix}\right\}$ when $k < p+1$ or if p is $\left\{\begin{matrix}\text{even}\\ \text{odd}\end{matrix}\right\}$ when $k > p$; *i. e.*, P_{0k} is $\left\{\begin{matrix}\text{on}\\ \text{not on}\end{matrix}\right\}$ the quadric when $\left\{\begin{matrix}k > p\\ k < p+1\end{matrix}\right\}$ whether p be odd or even. The above quadric can be denoted by

$$Q_{0,1,\cdots,p} = Q_{p+1,\cdots,2p+1} = R_y + \textstyle\sum_{i=1}^{i=p} y_i^2 = R_y + \sum_{i=p+1}^{i=2p+1} y_i^2$$

the notation indicating that it contains all the points P_{ik} of B for which i is drawn from $0, \cdots, p$ and k from $p+1, \cdots, 2p+1$. Again according to F. G., p. 257, R_y is an $\left\{\begin{matrix}E\\ O\end{matrix}\right\}$ quadric according as $p \equiv \left\{\begin{matrix}0, 3\\ 1, 2\end{matrix}\right\}$ mod. 2. The null point of $y_1 + \cdots + y_p = 0$ is on R_y if $\left\{\begin{matrix}\binom{p}{2}\\ \binom{p+1}{2}\end{matrix}\right\} \equiv 0$ mod. 2 when $p \equiv \left\{\begin{matrix}0\\ 1\end{matrix}\right\}$ mod. 2. Therefore in all cases $Q_{012\cdots p}$ is an E quadric. From its behavior with regard to the points of B we see that

THEOREM 3. *There exists for the basis configuration B a set of $\frac{1}{2}\binom{2p+2}{p+1} = \binom{2p+1}{p}$ median* E quadrics of the type $Q_{0,1,\cdots,p} =$*

* The term "median" refers to the fact that the number of squares is as close to $\frac{1}{2}(2p+2)$ as possible.

$Q_{p+1,\,p+2,\,\cdots\,2p+1} = R_y + y_1^2 + \cdots + y_p^2$ *where* R_y *is the quadric on none of the residual points of* B_0. *Each is determined by a separation of the* $2p + 2$ *bases into two sets of* $p + 1$ *each and contains the* $(p + 1)^2$ *points* P_{ik} *of* B *such that* B_i *and* B_k *are drawn from different sets. This set of quadrics is conjugate under the* $G_{(2p+2)!}$ *of* B.

Of the two ways of writing a quadric, the one contains an even, the other an odd, number of squares. Let us for convenience write it so that the number of squares is even or odd with $p + 1$. Then the quadric of Theorem 3 is $Q_{0,\,1,\,\cdots,\,p} = Q_{p+1,\,\cdots,\,2p+1} = R_y + y_{p+1}^2 + \cdots + y_{2p+1}^2$. Furthermore every quadric can be expressed as $Q' = Q_{0,\,1,\,\cdots,\,p} + \Sigma_{i=1}^{i=2k} z_i^2$ where the z's are any $2k$ of the y's and where $0 \leqq k < p + 1$. Suppose to fix ideas that l of the z's are $y_{p-l+1}, \cdots, y_p$ and $2k - l$ are $y_{p+1}, \cdots, y_{p+2k-l}$. Then $\Sigma_{i=1}^{i=2k} z_i^2$ is the null S_{2p-2} of the point $P_{p-l+1,\,\cdots,\,p+2k-l}$. This point is $\left\{\begin{matrix}\text{on}\\ \text{not on}\end{matrix}\right\}$ $Q_{0,\,1,\,\cdots,\,p}$ when $\binom{2k}{2} + l$ or when $k - l$ is $\left\{\begin{matrix}\text{even}\\ \text{odd}\end{matrix}\right\}$. Since we know when the point P_{ij} lies on $Q_{01\cdots p}$, and when it lies on Σz_i^2 it is easy to verify that the point lies on Q' if one of its subscripts is drawn from the set $0, 1, \cdots, p - l, p + 1, \cdots, p + 2k - l$ and the other from the set $p - l + 1, \cdots, p, p + 2k - l + 1, \cdots, 2p + 1$.

THEOREM 4. *Every quadric belonging to* C *is uniquely determined by a separation of the bases of* B *into two sets of* $p + 1 - 2k$ *and* $p + 1 + 2k$ *each. It contains those and only those points* P_{ij} *of* B *which belong to bases* B_i *and* B_j *drawn from different sets. It is an* $\left\{\begin{matrix}E\\ O\end{matrix}\right\}$ *quadric when* k *is* $\left\{\begin{matrix}\textit{even}\\ \textit{odd}\end{matrix}\right\}$. *The particular quadric associated with the separation* $B_0, \cdots, B_{p-2k}; B_{p+1-2k}, \cdots, B_{2p+1}$ *is* $Q_{0,\,\cdots,\,p-2k} = Q_{p+1-2k,\,\cdots,\,2p+1} = R_y + \Sigma_{i=p+1-2k}^{i=2p+1} y_i^2 = 0$. *If* $p + 1$ *is even we have for* $k = (p + 1)/2$ *an isolated quadric* $Q = Q_{0,\,1,\,\cdots,\,2p+1}$ *which contains none of the points of* B. *Under the* $G_{(2p+2)!}$ *of* B *the quadrics divide into* $\left\{\begin{matrix}(p + 1)/2\\ p/2\end{matrix}\right\}$ *conjugate sets if* p *is* $\left\{\begin{matrix}\textit{odd}\\ \textit{even}\end{matrix}\right\}$.

E. g., when $p = 2$ we have 10 E quadrics $Q_{ijk} = Q_{lmn}$ and 6 O quadrics $Q_i = Q_{jklmn}$. When $p = 3$ we have 35 E quadrics $Q_{ijkl} = Q_{mnop}$, 28 O quadrics $Q_{ij} = Q_{klmnop}$, and 1 E quadric $Q = Q_{01234567}$.

The question as to the incidence of a point P whose symbol has $2s$ subscripts with the quadric $Q_{0,\,\cdots,\,p-2k}$ is easily answered. Let the set of $2s$ subscripts of P be that set which does not contain 0 and of these let r be found in the set $0, \cdots, p - 2k$. Then the result of substituting in the equation of the quadric as given in Theorem 4 is $\binom{2s}{2} + 2s - r$ which is even or odd with $s + r$. But $s + r$ is even or odd with $s + (2s - r)$, or

with $(p + 1 - s) + (p + 1 + 2k - r)$, or with $(p + 1 - s) + (p + 1 - 2s - 2k + r)$ so that

THEOREM 5. *A point lies on a quadric if half the number of subscripts in a symbol for the point together with the number of subscripts common to this symbol and a symbol for the quadric is even.*

E. g., when $p = 3$ the E quadric $Q_{ijkl} = Q_{mnop}$ contains the 16 points $P_{im} = P_{jklnop}$, the 18 points $P_{ijmn} = P_{klop}$ and the point $P_{ijkl} = P_{mnop}$. The O quadric $Q_{ij} = Q_{klmnop}$ contains the 12 points $P_{ik} = P_{jlmnop}$ and the 15 points $P_{ijkl} = P_{mnop}$. The E quadric $Q = Q_{ijklmnop}$ contains the 35 points $P_{ijkl} = P_{mnop}$.

The result of adding to the quadric above the square of the null S_{2p-2} of the point P is evidently a quadric for which the $2s - r$ common squares reduce to zero mod. 2 so that the new quadric contains $r + (p + 1 + 2k) - (2s - r)$ squares. Hence

THEOREM 6. *The symbol for the quadric obtained by adding to a given quadric the square of the null S_{2p-2} of a given point has for its subscripts the aggregate of the subscripts in a symbol of the given quadric and a symbol of the given point less the subscripts common to the two given symbols.*

The group of the null system is generated by a conjugate set of involutions $I_{r_1, r_2, \ldots}$ each associated with a point $P_{r_1, r_2, \ldots}$. The involution $I_{r_1, r_2, \ldots}$ transforms the point $P_{s_1, s_2, \ldots}$ into itself if $P_{r_1, r_2, \ldots}$ and $P_{s_1, s_2, \ldots}$ are syzygetic; otherwise into the point $P_{r_1, r_2, \ldots} + P_{s_1, s_2, \ldots} = P_{r_1, r_2, \ldots; s_1, s_2, \ldots}$, where like subscripts cancel each other when adjoined. It is often convenient to effect these generating involutions upon the quadrics attached to the null system and this is accomplished as follows:

THEOREM 7. *The involution determined by a given point P with $2k$ subscripts transforms a given quadric whose symbol has σ subscripts in common with P, into itself when $k + \sigma$ is odd; but when $k + \sigma$ is even into a quadric whose symbol has for subscripts the sum of those of the given point and quadric.*

To prove this let $I_{r_1, r_2, \cdots, r_\sigma; s_1, s_2, \cdots, s_{2k-\sigma}}$ be the given involution and let the given quadric be

$$Q_{r_1, \cdots, r_\sigma; t_1, \cdots, t_{p+1-2l-\sigma}} = Q_{s_1, \cdots, s_{2k-\sigma}; u_1, \cdots, u_{p+1+\sigma+2l-2k}}.$$

Then Q contains the four sets of basis points

$$(a)\ P_{r,s},\ (b)\ P_{t,u},\ (c)\ P_{t,s},\ (d)\ P_{r,u},$$

which are transformed by I into respectively

$$(a)\ P_{r,s},\ (b)\ P_{t,u},\ (c)\ P_{t,\bar{r},\bar{s}'},\ (d)\ P_{u,\bar{r}',\bar{s}}.$$

Here a subscript r stands for any one of the r's, $\bar{r}$ for all the r's and $\bar{r}'$ for

all but one of the r's. Applying the rule of Theorem 5 we see that if $k + \sigma$ is odd the given quadric contains all four transformed sets, and that if $k + \sigma$ is even the quadric $Q_{s_1, \cdots, s_{2k-\sigma}; t_1, \cdots, t_{p+1-2l+\sigma}}$ contains all four transformed sets. Since there is one and only one quadric on the original four sets the transformed quadric on the transformed set is uniquely determined.

It seems that the above notation for the period and theta characteristics can hardly be improved. For not only is the notation (3) for the period characteristics and the notation of Theorem 4 for the theta characteristics each remarkably simple; but also the various fundamental conditions such as, (3) for the addition of half-periods, Theorem 1 for syzygetic or azygetic period characteristics, Theorem 4 for even or odd theta characteristics, Theorem 5 for the vanishing or non-vanishing of a given theta function for a given half-period, Theorem 6 for the theta function obtained by adding to the argument of a given theta function a given half-period, and Theorem 7 for the transformation of the theta functions, all are extremely easy to apply.

§ 2. An Isomorphism between the Median E Quadrics of B and the $(2p + 2)$-point in S_p.

The median E quadrics of Theorem 3 admit of a grouping into sets of $p + 2$ which appears as follows. Consider a $(p + 2, p)$ separation of the bases of B into a set $B_0, \cdots, B_{p+1}$ and a set $B_{p+2}, \cdots, B_{2p+1}$. If a base of the first set be added to the second a $(p + 1, p + 1)$ separation is obtained which determines a median E quadric. If a base of the second set be added to the first set, the $(p + 3, p)$ separation determines as in Theorem 4 a median O quadric. The $p + 2$ median E quadrics and the p median O quadrics so derived constitute a fundamental system ($F. S.$) of $2p + 2$ quadrics. In fact these quadrics can be gotten by beginning with $R_y + \Sigma_{i=1}^{i=p+1} y_i^2$ and adding in turn $y_1^2, \cdots, y_{2p+1}^2$. By referring to the table (F. G., p. 272) we see that in any $F. S.$ the number s of O quadrics is congruent to p mod. 4. When $s = p$ the values of k in the table are $p/2$; $(p + 2)/2$ or $(p - 1)/2$; $(p + 1)/2$ or $(p + 2)/2$; $p/2$ or $(p + 1)/2$; $(p - 1)/2$ according as $p \equiv 0, 1, 2, 3$ mod. 4. Hence both the E and the O quadrics in the $F. S.$ are median quadrics of the attached B.

Theorem 8. *The* $\frac{1}{2}\binom{2p+2}{p+1}$ *median E quadrics of a basis configuration B can be grouped in* $\binom{2p+2}{p}$ *ways into sets of $p + 2$ each such that each set can be supplemented by p median O quadrics to form a $F. S.$ of $2p + 2$ quadrics attached to B. Every $F. S.$ containing precisely p O quadrics can be obtained in this way.*

Consider now $2p+2$ points in a projective space S_p numbered 0, 1, $\cdots$, $2p+1$. One can form from them $\frac{1}{2}\binom{2p+2}{p+1}$ determinant products of the type $(0, 1, \cdots, p)(p+1, p+2, \cdots, 2p+1)$. This product corresponds in notation to the median E quadric $Q_{0,1,\cdots,p} = Q_{p+1,\cdots,2p+1}$. If we separate the points into a set of $p+2$ and a set of p, a determinant product can be formed by taking a point of the first set and adding it to the second set. Moreover the $p+2$ products thus formed satisfy a linear relation. Hence

THEOREM 9. *The* $\frac{1}{2}\binom{2p+2}{p+1}$ *determinant products that can be formed from* $2p+2$ *points in* S_p *are grouped* $p+2$ *at a time in* $\binom{2p+2}{p}$ *determinant identities precisely as the median E quadrics of Theorem* 8 *are grouped in F. S.*'s *containing p median O quadrics.*

For the case $p = 2$ it appears* that this formal isomorphism is vitalized by the existence of theta relations (*i. e.*, linear relations connecting the squares of the functions) which parallel the determinant identities and which lead to the determination of a set of 6 points in terms of theta modular functions. A similar set of relations is derived in § 4 for $p = 3$.

§ 3. The Determination of the Signs in Theta Relations.

In the usual notation for the theta function a proper half-period is given by a scheme of $2p$ numbers $\epsilon = \begin{pmatrix} \epsilon_1, & \cdots, & \epsilon_p \\ \epsilon_{p+1}, & \cdots, & \epsilon_{2p} \end{pmatrix}$ which are 0 or 1 but not all zero. This we have identified with the point x in the finite space S_{2p-1} such that $x_i = \epsilon_i$. The notation for the characteristic of an odd or even theta function of the first order is similarly $\eta = \begin{pmatrix} \eta_1, & \cdots, & \eta_p \\ \eta_{p+1}, & \cdots, & \eta_{2p} \end{pmatrix}$, $\eta_i = 0, 1$. This has been identified with the quadric in S_{2p-1}

(1) $\quad x_1x_{p+1} + \cdots + x_px_{2p} + \eta_1 x^2_{p+1} + \eta_{p+1}x_1^2 + \cdots + \eta_p x_{2p}^2 + \eta_{2p}x_p^2 = 0.$

The null system C, or invariant relation between two half-periods x and x', is then

$$x_1x'_{p+1} + x_{p+1}x_1' + \cdots + x_px_{2p}' + x_{2p}x_p' = 0.$$

In determining the coefficients of a theta relation one often adds a half-period to the argument. The change in a term is then given by the formula,†

$$\vartheta[\eta](u+\epsilon) = f(\epsilon) \cdot e^{-(\epsilon_1\eta_{p+1} + \cdots + \epsilon_p\eta_{2p})/4} \cdot \vartheta[\eta+\epsilon](u).$$

* Coble, "Point Sets and Allied Cremona Groups," Transactions of the American Mathematical Society, vol. XVI (1915), p. 155, § 9; referred to hereafter as P. S.

† Krazer, Lehrbuch der Thetafunktionen, p. 240 (VII).

Here $f(\epsilon)$, an exponential factor, depends only on the half-period ϵ and disappears from the relation; $\vartheta[\eta + \epsilon]$ corresponds in F. G. to the quadric obtained by adding to the quadric corresponding to $\vartheta[\eta]$ the square of the null S_{2p-2} of the point corresponding to ϵ, while the factor $e^{-(\Sigma \epsilon_i \eta_{p+i})/4}$ gives rise in any relation involving the squares of the thetas to ± 1 according as the congruence,

$$(2) \qquad \epsilon_1 \eta_{p+1} + \cdots + \epsilon_p \eta_{2p} \equiv 0 \text{ mod. } 2,$$

is or is not satisfied. But in general we are interested only in the ratio of two such signs say those determined by the terms $[\eta]$ and $[\eta']$. These are like or unlike according as

$$(3) \qquad K = \epsilon_1 \epsilon'_{p+1} + \cdots + \epsilon_p \epsilon_{2p}' \equiv 0 \text{ mod. } 2$$

is or is not satisfied when $(\epsilon') = [\eta] + [\eta']$. This relation (3) on the points ϵ, ϵ' of our F. G. is a correlation K which we wish to study in connection with the given null system C.

Let us first notice that the two spaces, S_{p-1}:

G_1, defined by $x_{p+1} = x_{p+2} = \cdots = x_{2p} = 0$, and

G_2, defined by $x_1 = x_2 = \cdots = x_p = 0$;

are such that the null S_{2p-2} of any point in the one contains it completely; i. e., each is a null space of C or a *Göpel space* as defined in F. G., p. 248. Points of G_2 taken as points ϵ, and points of G_1 taken as points ϵ', are singular points of K. But a point of G_1 taken as ϵ has a corresponding S_{2p-2}, the same under K as under C, which cuts G_2 in an S_{p-2}; and a point of G_2 taken as ϵ' has a corresponding S_{2p-2} under K or under C which cuts G_1 in an S_{p-2}. Hence

THEOREM 10. *The null system C determines a proper correlation between any two skew Göpel spaces. Conversely given a correlation between any two skew S_{p-1}'s in S_{2p-1}, this is determined by a unique null system C for which the S_{p-1}'s are Göpel spaces. The two Göpel spaces are contained in one and only one E quadric belonging to C.*

For if G_1, G_2, and K are given, a point x of G_1 determines under K an S_{p-2} in G_2 which is joined to G_1 itself by an S_{2p-2}, the null S_{2p-2} of x under C. The null S_{2p-2} of x' in G_2 under C is determined similarly. Any point not on G_1 or G_2 can be expressed in a single way as $x + x'$ whence its null S_{2p-2} under C is known and C is determined. Moreover it is clear from (1) that only the one quadric $[\eta] = 0$ contains G_1 and G_2.

If then there be isolated in the finite geometry a pair of skew Göpel spaces, the original E function $\vartheta(u)$ is isolated, the correlation K is determined, and thereby the signs in the theta relations are determined. Some

choices of G_1 and G_2 particularly convenient for the notation of § 1 will now be given.

In the space G_1 with coördinates $x_1, \cdots, x_p$ let the reference points in order and the unit point be denoted by $g_0, g_1, \cdots, g_p$; in G_2 with coordinates $x_{p+1}, \cdots, x_{2p}$ let the similar points be denoted by $\gamma_0, \gamma_1, \cdots, \gamma_p$. Then $\Sigma g = 0$ and $\Sigma\gamma = 0$. In the notation of § 1 consider the two sets of $p + 1$ points each:

$$(4) \qquad \begin{aligned} &P_{01}, P_{23}, P_{45}, \cdots, P_{2p,\,2p+1};\\ &P_{12}, P_{34}, P_{56}, \cdots, P_{2p+1,\,0}. \end{aligned}$$

The sum of the points in each set is identically zero so that each determines an S_{p-1}. Any two points of a set are syzygetic so that the S_{p-1}'s are Göpel spaces. No point of the one space can be a point of the other for if the even subscripts of two such points coincide the odd ones cannot. Hence the sets (4) can be identified with the sets g of G_1 and γ of G_2. The quadric corresponding to $\vartheta(u)$ is the median E quadric $Q_{0,\,2,\,4,\,\cdots,\,(2p)} = Q_{1,\,3,\,5,\,\cdots,\,(2p+1)}$.

The expression for any given point $P_{r_1,\,r_2,\,\cdots}$ in terms of the above points of G_1 and G_2 is quickly obtained by arranging $r_1, \cdots, r_{2k}$ in ascending order and filling the intervals between them with the intervening numbers each taken twice. E. g., $P_{1347} = P_{1223455667} = P_{12} + P_{23} + P_{45} + P_{56} + P_{67} = (P_{23} + P_{45} + P_{67}) + (P_{12} + P_{56}) = \begin{pmatrix} 011100\cdots \\ 101000\cdots \end{pmatrix}$. If two points P, P' are so expressed in the form $P = g + \gamma$, $P' = g' + \gamma'$ then P and P' satisfy K if g and γ' satisfy C.

When p is odd and the $2p + 2$ subscripts are divided into two sets ordered with regard to each other, say $\begin{pmatrix} 0, & 1, & \cdots, & p \\ p+1, & p+2, & \cdots, & 2p+1 \end{pmatrix}$, then the $p + 1$ points formed by taking p subscripts from the one set and the $(p + 1)$th complementary subscript from the other are such that their sum is identically zero and that any two are syzygetic whence they determine a G_1. The general point of G_1 is formed from j subscripts of the one line and $\left\{\begin{matrix} p + 1 - j \text{ complementary} \\ j \text{ corresponding} \end{matrix}\right\}$ subscripts of the other when j is $\left\{\begin{matrix} \text{odd} \\ \text{even} \end{matrix}\right\}$. By shifting the subscripts of the lower row forward one space a Göpel space G_2 is similarly determined which as before is skew to G_1. In this case the quadric containing G_1 and G_2 is the median E quadric $Q_{0,\,1,\,\cdots,\,p} = Q_{p+1,\,\cdots,\,2p+1}$ when $p \equiv 1 \pmod{4}$; but if $p \equiv 3 \pmod{4}$ it is $Q = Q_{0,\,1,\,\cdots,\,2p+1}$ (see (8)).

For the case $p = 2$ a pair of Göpel spaces (4) was used in P. S., p. 191.

For the case $p = 3$ a pair of Göpel spaces of the other sort is used in the next paragraph.

§ 4. Some Theta and Theta-modular Relations for $p = 3$.

We denote by $\pi_{r_1, r_2, \ldots}$ the half-period which corresponds to $P_{r_1, r_2, \ldots}$, and by $\vartheta_{r_1, r_2, \ldots}(u)$ the theta function which corresponds to $Q_{r_1, r_2, \ldots}$. Also let $\vartheta_{r_1, r_2, \ldots} = [\vartheta_{r_1, r_2, \ldots}(u)]_{u=0}$. The 5 E quadrics Q_{jklm}, Q_{iklm}, Q_{ijlm}, Q_{ijkm}, Q_{ijkl} together with the three O quadrics Q_{op}, Q_{pn}, Q_{no} lie in a *F. S.* (§ 2). Since there is a linear relation connecting any 9 theta squares, the 8 squares corresponding to the above quadrics together with $\vartheta^2(u)$ must be linearly related. The point P_{op} does not lie on any of the E quadrics nor on Q_{op} but does lie on Q_{pn} and Q_{no}. If then we substitute π_{op} for u in the assumed relation all the terms vanish except that containing $\vartheta_{op}(u)$ whence that term cannot occur. Similarly none of the odd thetas can occur and the form of the identity is

$$a\vartheta^2(u) - a_{jklm}\vartheta^2_{jklm}(u) - a_{iklm}\vartheta^2_{iklm}(u) - a_{ijlm}\vartheta^2_{ijlm}(u) - a_{ijkm}\vartheta^2_{ijkm}(u) - a_{ijkl}\vartheta^2_{ijkl}(u) = 0.$$

The point P_{jklm} is on Q and Q_{jklm} but is not on the other E quadrics whence for $u = \pi_{jklm}$ we get

$$a\vartheta^2(\pi_{jklm}) = a_{jklm}\vartheta^2_{jklm}(\pi_{jklm}).$$

From Theorem 6 we can set to within sign $a = \vartheta^2$, $a_{jklm} = \vartheta^2_{jklm}$. Reverting to (2) and (3) of § 3 if $\epsilon = \pi_{jklm} = \begin{pmatrix}\epsilon_1\epsilon_2\epsilon_3\\ \epsilon_4\epsilon_5\epsilon_6\end{pmatrix}$, then, since η = characteristic of $\vartheta(u) = \begin{pmatrix}0\,0\,0\\0\,0\,0\end{pmatrix}$, we see that $\eta_{jklm} = \begin{pmatrix}\epsilon_1\epsilon_2\epsilon_3\\ \epsilon_4\epsilon_5\epsilon_6\end{pmatrix}$. Thus $\eta + \eta_{jklm} = \epsilon' = \epsilon$ and the condition K in (3) § 3 is satisfied since P_{jklm} is on Q. Hence the signs of a and a_{jklm} are like and we see that

THEOREM 11. *After isolating the even theta function $\vartheta(u)$, the squares of the remaining 35 even thetas are linearly related with $\vartheta^2(u)$ by means of 56 identities of the type*

$$\vartheta^2\vartheta^2(u) - \vartheta^2_{jklm}\vartheta^2_{jklm}(u) - \vartheta^2_{iklm}\vartheta^2_{iklm}(u) - \vartheta^2_{ijlm}\vartheta^2_{ijlm}(u) - \vartheta^2_{ijkm}\vartheta^2_{ijkm}(u) - \vartheta^2_{ijkl}\vartheta^2_{ijkl}(u) = 0.$$

Let us denote by $\overline{\overline{ijkl\,mnop}}$ the determinant product formed from the coördinates of eight points $p_0, \cdots, p_7$ in S_3 when in each determinant the points are arranged in the natural order; and by $\epsilon_{ijklmnop}$ the sign of the permutation $ijklmnop$ after such an arrangement has been made. Now a typical pair of terms in a determinant identity is $\overline{ijkl}\,\overline{mnop} - \overline{ijkm}\,\overline{lnop}$.

Let ϵ_1 and $\epsilon_2 = -\epsilon_1$ be the signs of the permutations as written, and let a_1 and a_2 be respectively the number of inversions which are necessary to naturalize the order in the determinants of a product. Then $\epsilon_{ijklmnop} = (-1)^{a_1}\epsilon_1$ and $\epsilon_{ijklmnop} = (-1)^{a_2}\epsilon_2$. Hence the pair of terms

$$\epsilon_{ijklmnop}\,\overline{\overline{ijkl}\,\overline{mnop}} + \epsilon_{ijkmlnop}\,\overline{\overline{ijkm}\,\overline{lnop}}$$

is

$$(-1)^{a_1}\epsilon_1\,\overline{\overline{ijkl}\,\overline{mnop}} + (-1)^{a_2}\epsilon_2\,\overline{\overline{ijkm}\,\overline{lnop}}$$

or

$$(-1)^{2a_1}\epsilon_1\,\overline{ijkl}\,\overline{mnop} - (-1)^{2a_2}\epsilon_1\,\overline{ijkm}\,\overline{lnop}$$

or $\epsilon_1[\overline{ijkl}\,\overline{mnop} - \overline{ijkm}\,\overline{lnop}]$. Hence

THEOREM 12. *The 35 combinations of theta squares,* $\frac{1}{5}\vartheta^2\vartheta^2(u) - \vartheta^2_{ijkl}\vartheta^2_{ijkl}(u)$; *the 35 theta modular functions,* $\frac{1}{5}\vartheta^4 - \vartheta^4_{ijkl}$; *and the 35 determinant products formed from 8 points in S_3,* $\epsilon_{ijklmnop}\,\overline{\overline{ijkl}\,\overline{mnop}}$; *each satisfy the same set of 56 relations typified by*

$$[ijkl] + [ijkm] + [ijlm] + [iklm] + [jklm] = 0.$$

Since the 8 points in S_3 when self-associated, i. e., when they are the base points of a net of quadrics, depend upon 6 absolute constants, and the theta modular functions also depend upon 6 moduli it seems likely that the equations $\epsilon_{ijklmnop}\,\overline{\overline{ijkl}\,\overline{mnop}} = \frac{1}{5}\vartheta^4 - \vartheta^4_{ijkl}$ would serve for a parametric expression for 8 self-associated points in terms of the theta modular functions. Before this conclusion could be drawn however it would be necessary to show that the relations of higher degree satisfied by the determinant products (of which fourteen are linearly independent), and the relations (P. S., p. 165) which imply self-association, are likewise satisfied by the corresponding modular functions.

The 56 relations of Theorem 11 are associated with the basis configuration B which isolates $\vartheta(u)$. There being 36 B's we should expect to find 36.56 relations each occurring 6 times. There ought then to be 5.56 relations whose subscripts take another form. If we apply to B the involution I_{ijkn} the quadrics and points considered above are transformed into similarly situated quadrics and points. We can therefore infer that there is a linear identity connecting the six transformed even thetas which we shall prove is as follows:

$$\begin{aligned}(1)\quad &\vartheta^2_{ijkl}\vartheta^2_{ijkl}(u) + \vartheta^2_{ijkm}\vartheta^2_{ijkm}(u) + \vartheta^2_{ijkn}\vartheta^2_{ijkn}(u)\\ &= \vartheta^2_{ilmn}\vartheta^2_{ilmn}(u) + \vartheta^2_{jlmn}\vartheta^2_{jlmn}(u) + \vartheta^2_{klmn}\vartheta^2_{klmn}(u).\end{aligned}$$

We have only to verify the coefficients. If $u = \pi_{lm}$ all the terms except the first two vanish. Here $\eta = [ijkl]$, $\eta' = [ijkm]$, and $\epsilon' = \eta + \eta' =$

$[lm]$ while $\epsilon = [lm]$ also. Hence K in (3) § 3 is not satisfied since P_{lm} is not on Q. Then to within a common factor the terms become $\vartheta^2_{ijkl}\vartheta^2_{ijkm} - \vartheta^2_{ijkm}\vartheta^2_{ijkl}$ which checks the second coefficient. The third can be similarly checked. If $u = \pi_{jkmn}$ all terms vanish except the first and fourth. Then $\eta = [ijkl]$, $\eta' = [ilmn]$, $\epsilon' = \epsilon = [jkmn]$ and K is satisfied since P_{jkmn} is on Q. To within a factor the relation reduces to $\vartheta^2_{ijkl}\vartheta^2_{ilmn} = \vartheta^2_{ilmn}\vartheta^2_{ijkl}$ which checks the fourth coefficient. The fifth and sixth coefficients can be verified similarly. The relation (1) depends upon the separation of ijk, $lmn = lmn$, ijk from 8 things whence there are 56.5 of this form.

THEOREM 13. *The* 56.6 *theta relations for* $p = 3$ *which contain the five even functions which lie in any F. S. with three odd functions are comprised in the* 56 *relations of Theorem* 11 *and the* 56.5 *relations of type* (1).

If we denote the relation of Theorem 11 when transposed to the right by $\{ijklm\} = 0$ and (1) when transposed to the left by $\{ijk, lmn\} = 0$, then

$$\{ijkmn\} + \{ijknl\} + \{ijklm\} - \{lmnjk\} - \{lmnki\} - \{lmnij\} = 2\{ijk, lmn\}.$$

Thus none of the relations (1) are independent of those of Theorem 11. By means of the latter relations alone the 36 even theta squares can be expressed in terms of 15.

BALTIMORE, October 4, 1915.

ON CERTAIN REAL SOLUTIONS OF BABBAGE'S FUNCTIONAL EQUATION.*

By J. F. Ritt.

1. The first attempt to find the most general solution of the functional equation

$$f^n(x) = x,$$

where $f^2(x) = f[f(x)]$, and in general, $f^n(x) = f[f^{n-1}(x)]$, was made by C. Babbage, who published his results in Herschel's "Calculus of Finite Differences" in 1820. Babbage observed that if $f(x)$ is a particular solution of this equation, the formula

$$F(x) = \varphi^{-1}f\varphi(x),$$

$\varphi(x)$ being an arbitrary function, and $\varphi^{-1}(x)$ its inverse, will give an infinite number of other solutions. To verify this one need but calculate $F^n(x)$. For the particular solution $f(x)$ one can use, for instance, the well known "periodic transformation"

$$f(x) = \frac{\alpha + \beta x}{\gamma + \delta x}$$

where

$$\delta = -\frac{\beta^2 - 2\beta\gamma\cos\dfrac{2k\pi}{n} + \gamma^2}{2\alpha\left(1 + \cos\dfrac{2k\pi}{n}\right)},$$

k being any integer prime to n.†

To avoid the very ambiguous term "periodic" we shall refer briefly to a solution of Babbage's equation as a "function of order n," understanding therein that n is the lowest order of the function. Also, we may occasionally say "$f(x)$ circulates with a period n."

Babbage called $F(x) = \varphi^{-1}f\varphi(x)$ the general solution of his equation, but, as Pincherle explains,‡ the term "general" refers only to the presence of the arbitrary function $\varphi(x)$ and does not signify complete generality. In fact, Babbage§ himself gave an example to show that his solution

* Read under a different title before the American Mathematical Society, February 27, 1915.

† Boole, "Calculus of Finite Differences," p. 208.

‡ Encyclopédie des Sciences Mathématiques, II, 26.

§ Philosophical Transactions, 1817, p. 206.

lacked generality, but his conclusions have little value from the point of view of the modern theory of functions.

Babbage's equation was also studied by Leau,* who confined himself to the analytic solutions.†

The object of this paper is to study a certain class of real solutions of Babbage's equation, and to determine whether they are all contained in the formula $\varphi^{-1}f\varphi(x)$. We shall, with certain assumptions as to $\varphi(x)$, $\varphi^{-1}(x)$ and the solutions of the equation, find the most general solution to be, when $n > 2$,

$$F(x) = \varphi^{-1}f^p\varphi(x)$$

where p takes on all integral values prime to n. It will be necessary to consider separately the case of $n = 2$.

2. It is evident that $\varphi^{-1}f\varphi(x)$ will be the most general solution if, when given two solutions, $F(x)$ and $f(x)$, we can find a function $\varphi(x)$ such that

$$F(x) = \varphi^{-1}f\varphi(x).\ddagger$$

We shall try to construct such a function $\varphi(x)$. Suppose $\varphi(a) = b$. Observing that if $\varphi(x)$ exists, $F^p(x) = \varphi^{-1}f^p\varphi(x)$, we must have

$$\varphi F^p(a) = f^p\varphi(a) = f^p(b).$$

In particular, $\varphi F^n(a) = f^n(b)$, and this condition reduces to $\varphi(a) = b$, as assumed. Conversely, if we define $\varphi(x)$ as indicated for the n points $x = F^p(a)$, $(p = 0, 1, 2, \cdots, n - 1)$, we shall have, for those points, $F(x) = \varphi^{-1}f\varphi(x)$. Similarly, if we put $\varphi(a') = b'$, we can define $\varphi(x)$ for n more points. If some of the points $F^p(a)$, $F^p(a')$ coincide, it may be necessary to take $\varphi(x)$ as a many-valued function.

Thus we can proceed, defining $\varphi(x)$ for an indefinite number of points, and the generality of Babbage's solution appears quite complete. However, if we propose to define $\varphi(x)$ for the entire domain of x and to make it one-valued and generally continuous, the problem is far from being settled. It might be proposed to define $\varphi(x)$ arbitrarily in some interval (a, a'), say as some continuous function $\Phi(x)$, then to define it as $f\Phi F^{-1}(x)$ in the interval $[F(a), F(a')]$, as $f^2\Phi F^{-2}(x)$ in $[F^2(a), F^2(a')]$, etc., according to the principle explained above. But one might object, for instance, that however small the interval (a, a') be taken, $F(x)$ will lie in (a, a') for some values of x in (a, a'), and that it will then be out of the question to define $\varphi(x)$ arbitrarily.

Thus the problem will be solved if we can divide the domain of x into n intervals, such that if x lies in the first interval, $f(x)$ will lie in the second,

* Bulletin de la Société Mathématique de France, 1898.

† See also A. A. Bennet, Annals of Mathematics, Vol. 17, 1915, p. 37.

‡ This principle was used by Babbage, *Phil. Trans.*, 1817.

$f^2(x)$ in the third, etc. It will then be possible to take $\varphi(x)$ arbitrarily in one of these intervals, and its behavior in the entire domain of x will be determined.

3. In order to advance in this direction, we take up the study of a function $f(x)$, of order n, making the following assumptions.

(1) It is defined uniquely for all real values of x, including $+\infty$ and $-\infty$.

(2) It is continuous everywhere, except at an isolated set of points, where it becomes infinite. At these it may approach the same infinity from both sides, in which case it is to be defined as that infinity, or it may approach $+\infty$ from one direction and $-\infty$ from the other, in which case it is to be defined either as $+\infty$ or as $-\infty$. Should $f(x)$ approach a limit, finite or infinite, as x approaches $+\infty$, it is to be defined, for $x = +\infty$, as that limit. If no limit is approached, the definition is to be made in some other manner. Similar remarks apply to $x = -\infty$. If $f(+\infty) = f(-\infty)$, we shall not distinguish between $+\infty$ and $-\infty$.

(3) $f^n(x) = x$ for every value of x, but the integer p not being divisible by n, we can find, in any interval, however small, a point x_1 such that $f^p(x_1) \neq x_1$.

The first two of the above assumptions are to apply also to $\varphi(x)$ and to $\varphi^{-1}(x)$.

4. The function $f^p(x)$, p being any positive integer, takes on every real value for some value of x. For, taking the positive integer k such that $kn > p$,

$$f^p[f^{kn-p}(c)] = f^{kn}(c) = c,$$

where c is any real number, including $\pm\infty$.

Also, the value c is assumed for only one value of x.* For if

$$f^p(\alpha_1) = f^p(\alpha_2) = c,$$

then

$$f^{kn}(\alpha_1) = f^{kn-p}(c) = f^{kn}(\alpha_2)$$

or

$$\alpha_1 = \alpha_2.$$

Thus $f^p(x)$, which is itself one-valued by § 3, has a one-valued inverse, defined for all real values of x, including $+\infty$ and $-\infty$. It follows without difficulty that if we denote the inverse of $f^p(x)$ by $f^{-p}(x)$, we may, without ambiguity, allow the index p to assume all integral values, positive and negative, using the formula $f^p[f^q(x)] = f^{p+q}(x)$ and calling $f^p(x)$ and $f^q(x)$ identical when $p \equiv q$ [mod. n].

5. From the fact that $f(x)$ and its inverse are both one-valued, it follows that $f(x)$ increases or decreases monotonically in any interval in

* This fact was first observed by Leau, loc. cit.

which it is continuous. For the same reason, if $f(x)$ becomes discontinuous at a point it must approach $+\infty$ from one direction and $-\infty$ from the other. Finally, if $f(x)$ is discontinuous for $x = c$, it will be bounded, and therefore continuous, in any interval not containing c. In fact $f(x)$ must be bounded in the entire domain of x, if we reject some interval about c. The formal proof of these facts will be found very simple.

6. From §§ 4 and 5, we see that the following are the only possible types of $f(x)$:

Type I. Discontinuous at one point c, and monotone increasing wherever continuous.

As x increases from c to $+\infty$, $f(x)$ increases monotonically to a finite limit a, taking on all values less than a. As x decreases from c to $-\infty$, $f(x)$ decreases monotonically from $+\infty$, and since, by § 4, it must now take on all values greater than a, and these only, it must approach a as x approaches $-\infty$. That is, the graph of $f(x)$ consists of two infinite branches, with the asymptotes $x = c$, and $y = a$.

If $x_1 < x_2$ and c does not lie in the interval (x_1, x_2),

$$f(x_1) < f(x_2),$$

but if c does lie in the interval (x_1, x_2),

$$f(x_1) > f(x_2).$$

Still we shall call $f(x)$ *increasing.*

Type II. Discontinuous at one point c, and monotone decreasing wherever continuous.

Similar remarks apply as to *Type I.*

Type III. Continuous everywhere, and monotone decreasing.

As x increases from $-\infty$ to $+\infty$, $f(x)$ decreases monotonically from $+\infty$ to $-\infty$.

Type IV. Continuous everywhere, and monotone increasing.

As x increases from $-\infty$ to $+\infty$, $f(x)$ increases monotonically from $-\infty$ to $+\infty$.

Since, in the discussion of $f(x)$, we have used the fact that $f^n(x) = x$ only to show that $f^{-1}(x)$ satisfies the first assumption of § 3, it is clear that $\varphi(x)$ and $\varphi^{-1}(x)$ also belong to the above four types.*

---

* Dr. G. M. Green has made the interesting observation that the four types of functions under consideration are all contained in the formula

$$f(x) = \frac{a + bF(x)}{c + dF(x)}$$

where a, b, c and d are real constants such that $ad - bc \neq 0$, and where $F(x)$ is of *Type IV.*

It is not difficult to see that $f(x)$ as thus determined will always be of one of our four types.

7. We leave to the reader the formal work of showing that a function formed by compounding any number of functions of the four types of § 6 will belong to one of those four types, and will be increasing or decreasing according as an even number or an odd number of decreasing functions are employed.

In particular, if $f(x)$ circulates with a period n, every $f^p(x)$ will belong to one of the four types of § 6. If p is even, $f^p(x)$ will be increasing, and if p is odd, $f^p(x)$ will have the same monotonic character as $f(x)$.* Also, if d is the highest common divisor of p and n, then $f^p(x)$ circulates with a period n/d, and m being any integer, not divisible by n/d, we can find in any interval, however small, a point x_1 such that $f^{mp}(x_1) \neq x_1$.

8. Consider a function $f(x)$ of order $n > 1$, and of *Type I*, discontinuous for $x = c$. Suppose $f(a) = a$ for some finite a. Then $f^p(a) = a$ for every p, so that every $f^p(x)$ is finite, and therefore continuous, at a. Thus we can find a δ, where $0 < \delta < |c - a|$, such that $|f^p(x) - a| < |c - a|$ for $|x - a| < \delta$ and for every p. Take a point x_1 in the interval $(a - \delta, a + \delta)$ such that $f(x_1) \neq x_1$. This is possible by the third assumption of § 3. Then c cannot lie between $f^p(x_1)$ and $f^q(x_1)$ for any values of p and q, else either $|f^p(x_1) - a|$ or $|f^q(x_1) - a|$ would be greater than $|c - a|$. Hence, since $f(x)$ is monotone increasing in any interval not containing c,

$$f^{p+1}(x_1) > f^{q+1}(x_1), \text{ if } f^p(x_1) > f^q(x_1).$$

To fix our ideas, suppose $f(x_1) > x_1$. Then

$$x_1 < f(x_1) < f^2(x_1) < \cdots < f^n(x_1) = x_1.$$

Hence we cannot have $f(a) = a$ for any finite a.

Suppose now $f(x)$, of order $n > 1$, is of *Type IV*. Take a point x_1 such that $f(x_1) \neq x_1$; suppose $f(x_1) > x_1$, for instance. Then, as above,

$$x_1 < f(x_1) < f^2(x_1) < \cdots < f^n(x_1) = x_1,$$

from which it is evident that $f(x)$ cannot be of *Type IV* except in the trivial case of $f(x) = x$.

We see now that if $f(x)$, of order $n > 1$, is of *Type I*, $f^p(x)$ will be of *Type I* if p is not divisible by n; for $f^p(x)$ is increasing and, being of order greater than unity, cannot be of *Type IV*. Then we cannot have

Also it is evident that the formula gives all functions of *Type III* and *Type IV*. If $f(x)$ is of one of the first two types, it is easy to see that $F(x) = f(x) / [f(x) - f(\infty)]$ is of *Type III* or of *Type IV*. Solving this equation for $f(x)$, we find $f(x)$ exhibited in the desired form. We inserted this note while correcting the proof, but see no way, as yet, for putting this interesting result to profit, principally because it does not lead to a simple expression for $f^p(x)$.

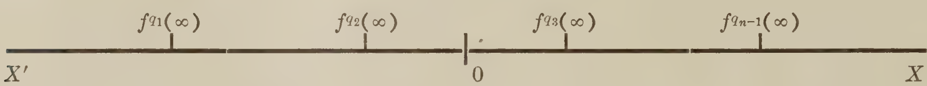
* It follows at once that if n is odd $f(x)$ must be increasing, else $f^n(x) = x$ must decrease as x increases.

$f^p(a) = a$ for any finite a, if p is not divisible by n. Also, if $f^p(\infty) = \infty$, then $f^p(c) = c$, so that the same applies to ∞.

9. The reader will find no difficulty in showing that if $f(x)$ is of *Type II*, we must have $f(x) = x$ for two points, one in the interval $(-\infty, c)$, the other in $(c, +\infty)$. But if $f(x)$ is of order $n > 2$, $f^2(x)$ will be of order greater than unity, and will be of *Type I*. Then since $f^2(x) = x$ when $f(x) = x$, we are led to a contradiction of the results of the preceding article. Hence, *Type II* is impossible for $n > 2$.

If $f(x)$, of order $n > 2$ is of *Type III*, then $f^2(x)$ will be of order greater than unity and of *Type IV*. Hence *Type III* is impossible for $n > 2$.

10. Thus, n being greater than 2, and p not divisible by n, every $f^p(x)$ will be of *Type I* and will be discontinuous for $x = f^{1-p}(c)$. These points of discontinuity are all distinct for $p = 1, 2, \cdots, n-1$, else we would have $f^{q_1}(c) = f^{q_2}(c)$, or $f^{q_1-q_2}(c) = c$, where $q_1 \neq q_2$. We shall plot these points, designating the point of discontinuity of $f^p(x)$ by $f^{-p}(\infty)$, and using the indices $q_1, q_2, \cdots, q_n$, of which none is divisible by n, and of which no two differ by a multiple of n.



The function $f^{-q_1}(x)$ is continuous and monotone increasing in any interval not containing the point $f^{q_1}(\infty)$. Hence, for values of r from 2 to $n-2$ inclusive, we must have

$$f^{-q_1}[f^{q_{r+1}}(\infty)] > f^{-q_1}[f^{q_r}(\infty)].$$

In virtue of this inequality we must have

$$f^{-q_1}[f^{q_{r+1}}(\infty)] = f^{q_r}(\infty),$$

and consequently

$$q_{r+1} \equiv q_r + q_1 \quad [\text{mod. } n],$$

where, now, r may also be unity.

Therefore, neglecting multiples of n,

$$q_r = rq_1,$$

for all values of r from 1 to $n-1$ inclusive.

Since no two of the indices, $q_1, q_2, \cdots, q_{n-1}$, are congruent to each other with respect to n, the index q_1 must be prime to n. We shall now designate q_1 by the letter m, which we shall call the *scale* of $f(x)$.

It is clear, now, that as x increases from $-\infty$ to $f^m(\infty)$, $f^m(x)$ will increase continuously from $f^m(\infty)$ to $f^{2m}(\infty)$, $f^{2m}(x)$ similarly from $f^{2m}(\infty)$ to $f^{3m}(\infty)$ and finally $f^{(n-1)m}(x)$ will increase monotonically from $f^{(n-1)m}(\infty)$

to $+\infty$. Hence we have accomplished the main purpose of our investigation, for $n > 2$, having divided the domain of x into n intervals such that if x_1 lies in one of them, $f(x_1)$ will lie in a second, etc. If we arrange the numbers $f^p(x_1)$ in order of magnitude (ascending), thus:

$$f^{q_1}(x_1),\ f^{q_2}(x_1),\ \cdots,\ f^{q_n}(x_1),$$

the point $f^{q_p}(x_1)$ will lie between $f^{(p-1)m}(\infty)$ and $f^{pm}(\infty)$. The indices $q_1, q_2, \cdots, q_n$, will form an arithmetic progression whose constant difference is the scale of $f(x)$; that is, neglecting multiples of n.

11. If $f(x)$ is of order n, $f^p(x)$ will have the same order if p and n are relatively prime. Also, if $f(x)$ has the scale m and $f^p(x)$ the scale m',

$$m \equiv pm' \quad [\text{mod. } n].$$

If, then, we can find a function $f(x)$ of order n, we can determine p through the above congruence so that $f^p(x)$ has any desired scale, prime to n. The "periodic transformations" assure us of the existence of functions of all orders, so that all scales prime to n are possible.

12. The reader will find it interesting to verify the results of the foregoing sections in the case of the linear fractional "periodic transformations." For instance, the increasing character of $f(x)$, for $n > 2$, can be shown by differentiating $f(x)$ and using the formula for δ of § 1. Using formulæ for $f^p(x)$ as given by Boole, it is seen that the roots of the equation $f^p(x) = x$ are imaginary unless p is a multiple of n. Also, a little investigation will reveal the scale of $f(x)$. One will find

$$km \equiv \pm 1 \quad [\text{mod. } n],$$

where k is the integer, prime to n, mentioned in § 1, and where the second member has the same sign as

$$\frac{\beta + \gamma}{2\delta} \sin \frac{2k\pi}{n}.$$

13. The argument of § 9 which excludes *Types II* and *III* for $n > 2$, does not apply when $n = 2$. In fact, when $n = 2$, *Types I, II* and *III* exist, as the following examples show.

Type I.

$$f(x) = \frac{\alpha - \beta x}{\beta + \delta x}, \qquad \text{where} \qquad \beta^2 + \alpha\delta < 0.$$

Type II.

$$f(x) = \frac{\alpha - \beta x}{\beta + \delta x}, \qquad \text{where} \qquad \beta^2 + \alpha\delta > 0.$$

Another example is $f(x) = \log(1 - e^x)$ in $(-\infty, 0)$ and $f(x) = -\log(1 - e^{-x})$ in $(0, +\infty)$.

Observe that $f(x) = x$ for two real values of x.

Type III.

$$f(x) = a - x.$$

Observe that $f(x) = x$ for a single value of x.

In the case of $n = 2$, the graph of $f(x)$ must be symmetric with respect to the line $y = x$.

14. We shall now subdivide the domain of x for functions of order 2. It is evidently necessary to consider only *Types II* and *III*.

(*a*) Suppose $f(x)$ is of *Type II*. Let c be the point of discontinuity, x_1 and x_1' the points for which $f(x) = x$, so that $x_1 < c < x_1'$.

The domain of x is divided as desired by the points x_1 and x_1'. As x increases from x_1 to x_1', $f(x)$ decreases monotonically from x_1, and passing through $-\infty$ and $+\infty$, decreases again to x_1'.

(*b*) When $f(x)$ is of *Type III*, the domain of x is divided as desired by the point for which $f(x) = x$.

15. We return now to our original problem; that is, given two functions, $f(x)$ and $F(x)$, both of order n, we shall see when there exists a function $\varphi(x)$, which, together with its inverse, satisfies the first two assumptions of § 3, and such that

$$F(x) = \varphi^{-1} f \varphi(x).$$

It will be necessary to consider separately the case of $n = 2$.

Thus, taking $n > 2$, we suppose first that $F(x)$ and $f(x)$ have the same scale m.

Take

$$\varphi F^{pm}(\infty) = f^{pm}(\infty) \qquad (p = 1, 2, \cdots, n-1),$$

$$\varphi(+\infty) = +\infty, \qquad \varphi(-\infty) = -\infty.$$

Take $\varphi(x)$ as any continuous monotone function in the interval $[F^m(\infty), F^{2m}(\infty)]$. Then, following the method of § 2, we can determine $\varphi(x)$ for all values of x. We will find $\varphi(x)$ to be of *Type IV*.

If $F(x)$ has the scale m and $f(x)$ the scale $n - m$, we can take $\varphi F^{pm}(\infty) = f^{pm}(\infty)$, $\varphi(+\infty) = -\infty$ and $\varphi(-\infty) = +\infty$. Proceeding as above, we can find a function $\varphi(x)$ of *Type III*.

Now the reader will not find it very difficult to show that if $f(x)$ has the scale m and if $\varphi(x)$ belongs to one of the four types of § 6, the scale of $F(x) = \varphi^{-1} f \varphi(x)$ will be m or $n - m$ according as φ is increasing or decreasing. That is, neglecting multiples of n, the square of the scale is invariant under the $\varphi^{-1} - \varphi$ transformation.* It is impossible, then, to

* To this numerical invariant corresponds an algebraic invariant of all linear fractional functions if we take $\varphi(x)$ linear fractional.

construct $\varphi(x)$ as desired except in the two cases above. Thus Babbage's formula lacks generality unless we give more liberty to $\varphi(x)$.

However, in the general case, where $F(x)$ is of scale m and $f(x)$ of scale m', we can, by § 11, determine the integer p, prime to n, so that $f^p(x)$ is of scale m. Then we have

$$F(x) = \varphi^{-1}f^p\varphi(x),$$

where $\varphi(x)$ is of *Type IV*. This formula, then, will give all solutions of Babbage's equation which satisfy the assumptions of § 3, in terms of one such solution. Observe that while we get all solutions by restricting $\varphi(x)$ to *Type IV*, we can also use any of the other types. In fact it would be easy to show that we can replace the $\varphi(x)$ of *Type IV* by another function of *Type I*, discontinuous at any desired point. Observe also that $\varphi(x)$, except for its monotone and continuous nature, is arbitrary within the entire interval $[F^m(\infty), F^{2m}(\infty)]$.

If we give to p values which are not prime to n, the order of $F(x)$ will be a divisor of n. It is easily seen that our formula gives all functions whose orders divide n, except perhaps those of *Types II* and *III* of order 2, should n be even. It will be seen that we do not get these.

16. When $n = 2$ it is necessary to take account of the types to which $f(x)$ and $F(x)$ belong.

Case I. $f(x)$ and $F(x)$ both of *Type I.*

In this case the method of § 15 applies and $\varphi(x)$ can be found, continuous everywhere.

Case II. $f(x)$ and $F(x)$ both of *Type II.*

Let x_1 and x_1' be the points for which $F(x) = x$, X_1 and X_1' the points for which $f(x) = x$. Let c be the point of discontinuity of $F(x)$, and C that of $f(x)$.

Take $\varphi(x_1) = X_1$, $\varphi(c) = C$, $\varphi(x_1') = X_1'$. Take $\varphi(x)$ monotone and continuous in (x_1, x_1'). Then $\varphi(x)$ will be monotone and continuous throughout the domain of x.

Case III. $f(x)$ and $F(x)$ both of *Type III.*

Let $F(x_1) = x_1$ and $f(X_1) = X_1$. Take $\varphi(x_1) = X_1$, and $\varphi(x)$ continuous, and monotone increasing to $+\infty$ or monotone decreasing to $-\infty$ in $(x_1, +\infty)$. Then $\varphi(x)$ will be continuous and monotone everywhere.

Case IV. $f(x)$ of *Type I* and $F(x)$ of *Type II.*

Evidently we cannot determine $\varphi(x)$ as desired, for by § 7, $\varphi^{-1}f\varphi(x)$ has the same monotonic character as $f(x)$.

Case V. $f(x)$ of *Type I* and $F(x)$ of *Type III.*

As in *Case IV* we cannot determine $\varphi(x)$ as desired.

Case VI. $f(x)$ of *Type II* and $F(x)$ of *Type III.*

Let x_1 be the point for which $F(x) = x$, X_1 and X_1' the points for which $f(x) = x$. Take $\varphi(x_1) = X_1$ and $\varphi(\pm \infty) = X_1'$ and $\varphi(x)$ monotone and continuous in the interval $(x_1, +\infty)$. Then $\varphi(x)$ will result monotone wherever continuous, with a single discontinuity. As an example, if $f(x) = 1/x$ and $F(x) = -x$ we can take $\varphi(x) = (x-1)/(x+1)$. A little consideration will show that it is impossible to avoid the discontinuity of $\varphi(x)$ and that the principle employed above is the only one which will give $\varphi(x)$ as desired.

Summarizing the results of this section, we see that in the case of $n = 2$, the formula

$$F(x) = \varphi^{-1}f\varphi(x)$$

gives all solutions of *Type I*, and only those, if $f(x)$ is of *Type I*, and gives all solutions of *Types II* and *III*, and only those, if $f(x)$ is either of *Type II* or *Type III*.

17. Let $f(x)$ and $F(x)$ be two differentiable functions of order 2.

Let

$$\varphi(x) = \log \frac{d}{dx} f(x) \quad \text{and} \quad \Phi(x) = \log \frac{d}{dx} F(x).$$

Put

$$y = f(x) \quad \text{and} \quad x = f(y).$$

Then

$$1 = \frac{d}{dx} f(y) = \frac{d}{dy} f(y) \frac{d}{dx} f(x)$$

or

$$\varphi(y) + \varphi(x) = 0,$$

so that

$$y = f(x) = \varphi^{-1}[-\varphi(x)],$$

or

$$\varphi f \varphi^{-1}(x) = -x.$$

Similarly,

$$\Phi F \Phi^{-1}(x) = -x,$$

so that

$$F(x) = \Phi^{-1}\varphi f \varphi^{-1}\Phi(x),$$

and putting

$$\varphi^{-1}\Phi(x) = \psi(x),$$

we have

$$F(x) = \psi^{-1}f\psi(x).$$

This reduction can be effected,* irrespective of the types of $F(x)$ and $f(x)$, but, as a general rule, $\psi(x)$ and $\psi^{-1}(x)$ will not conform with the first two assumptions of § 3.

Columbia University.

* If $f(x) = a - x$ we get an identity on differentiating, but as we have, in this case,

$$f(x) - \frac{a}{2} = -\left(x - \frac{a}{2}\right),$$

the reduction can proceed.

NOTE ON THE PRECEDING PAPER.

By Albert A. Bennett.

If we be given a one-valued real function $F(x)$, defined for all real values of x including infinity, and if $F^n(x) \equiv x$, where the index n refers to iteration, then we see that without imposing any further restriction, $F^{-1}(x)$, $[\equiv F^{n-1}(x)]$, is a one-valued real function defined for all real values of x including infinity. If we look upon $x' = F(x)$ as a transformation of a set of real numbers into itself, it will be a one-one reciprocal transformation. We may take a circle and choose upon it an arbitrary point to represent infinity, and then represent all the finite real points, preserving order, by the remaining points of the circle: for example, one such method of representation is to let $x = \tan \frac{1}{2}\theta$, where θ is the central angle. The transformation $x' = F(x)$ then determines a one-one reciprocal transformation among the points of the circle. An obvious restriction is to require this transformation to be also continuous. *Thus, the problem of Mr. Ritt's paper may be viewed as that of discussing in terms of Analysis Situs, the different types of continuous one-one reciprocal transformations of period n of a circle into itself.*

By transforming through a continuous, sense-preserving, one-one reciprocal transformation, the above problem of Analysis Situs becomes *normalized* into a problem of the periodic Euclidean transformations of period n of a circle into itself, such as is discussed, for instance, in any treatise on linear fractional transformations in the complex plane. A mere statement of cases will here suffice:

1. Reflections (altering sense)—period two.
   (*a*) Leaving "∞" and another point fixed.
   (*b*) Leaving two points, but not "∞", fixed.
2. Rotations (preserving sense)—period n.
   Leaving no point fixed—$(n - 1)$ cases; rotations of $1/n, 2/n, 3/n, \cdots, (n - 1)/n$ circumferences.

In terms of *functions*, $y = F(x)$, these become:

1. Monotonic decreasing—period two.
   (*a*) No finite vertical or horizontal asymptote. Curve cuts line $y = x$ in one finite point.
   (*b*) A finite vertical and a finite horizontal asymptote. Curve cuts $y = x$ in two finite points.
2. Monotonic increasing—period n.

A finite horizontal and a finite vertical asymptote. Curve does not cut $y = x$, — $(n - 1)$ cases, corresponding to $f, f^2, f^3, \cdots, f^{n-1}$.

Princeton, N. J.,
September, 1915.

AN ELEMENTARY EXPOSITION OF THE THEORY OF THE GAMMA FUNCTION.

By J. L. W. V. Jensen.

Authorized translation from the Danish by T. H. Gronwall.[1]

Although there exists a very considerable literature concerning the Gamma function, it is but recently that a connected exposition of the properties of this function has appeared, with a fairly complete bibliography[2, 3]. However, the monograph referred to does not seem to lay any stress on a rigorous exposition, and the treatment is neither carried out from any definite point of view, nor based on the simplest foundations. The present paper may therefore not be entirely devoid of interest, since it gives a short, but rigorous and fairly complete theory of the Gamma and allied functions, based on the elementary theory of functions, i. e., that part of the theory of infinite series and products, and particularly power series, which may be treated in a simple and natural way without the aid of the calculus. Thus I shall deal below with the properties of certain functions of a complex variable, while the connection with definite integrals must fall beyond the scope of the present paper; the transition to these may however be made with the utmost facility at almost any stage, as I may perhaps show on another occasion.[4]

1. Definitions. When we attempt to determine a single-valued function $f(s)$ satisfying the functional equation

$$f(s+1) = sf(s), \tag{1}$$

it is necessary to inquire first if this problem has more than one solution.

[1] Gammafunktionens Theori i elementær Fremstilling. Nyt Tidsskrift for Mathematik, Afdeling B, vol. 2 (1891), pp. 33–35, 57–72, 83–84.

It is hoped that a translation of this monograph, which combines the merits of conciseness and elegance in an unusual degree, will be of general interest. The footnotes enclosed in square brackets [ ] have been added by the translator, and contain additional references to the literature, some propositions supplementary to the text, and finally explanations, in the infrequent cases where such have been thought advisable.

[2] Brunel, G. Monographie de la fonction Gamma. Mémoires de la société des sciences physiques et naturelles de Bordeaux, ser. 3, vol. 3 (1886), pp. 1–184.

[3] [Of the literature published since the appearance of the original, we shall only mention here Brunel's article on definite integrals in Encyclopedie der Math. Wissensch. and particularly Nielsen, N., Handbuch der Theorie der Gammafunktion, Leipzig, Teubner, 1906. The latter contains a very complete bibliography.]

[4] [This project was unfortunately never carried out.]

Let $F(s)$ denote, for the moment, some definite and single-valued solution, and write $f(s) = p(s) \cdot F(s)$; it is then seen at once that the relation

$$p(s+1) = p(s)$$

constitutes the necessary and sufficient condition that $f(s)$ shall satisfy (1). In other words, whenever we have found one solution of the functional equation, there exists an infinity, since any solution retains its character as such upon multiplication by an arbitrary periodic function having the additive period unity, and conversely it is evident that every possible solution is obtainable in this manner.

We[5] are led to the discovery of a particular solution of (1) by the following considerations.

For any particular solution $F(s)$ that may exist, we have, by (1),

$$(2) \qquad \begin{aligned} F(s) &= \frac{F(s+1)}{s} = \frac{F(s+2)}{s(s+1)} = \cdots \\ &= \frac{F(s+n)}{s(s+1)\cdots(s+n-1)} \end{aligned}$$

where n, as throughout the following discussion, denotes a positive integer, and s a variable which never equals a negative integer or zero but may take any other value real or *complex*.

Moreover when s is a positive integer, we have

$$F(s+n) = (s+n-1)(s+n-2)\cdots n \cdot (n-1)!F(1),$$

and therefore

$$(a) \qquad \lim_{n\to\infty} \frac{F(s+n)}{(n-1)!n^s} = F(1).$$

Therefore, since

$$(b) \qquad \frac{F(s+n)}{s(s+1)\cdots(s+n-1)} = \frac{F(s+n)}{(n-1)!n^s} \cdot \frac{(n-1)!n^s}{s(s+1)\cdots(s+n-1)},$$

we have, at least for positive integral values of s,

$$(c) \qquad F(s) = F(1) \cdot \lim_{n\to\infty} \frac{(n-1)!n^s}{s(s+1)\cdots(s+n-1)}.$$

But, as we proceed to show, the second factor on the right has a definite value for every finite value of s, real or complex, except $0, -1, -2, \cdots$; it being understood that here and in what follows n^s is the num-

[5] The discussion from this point to that similarly marked below replaces the corresponding discussion in Jensen's paper. The proof of the uniform convergence of the infinite product which defines $\Gamma(s)$, however, is that given by Jensen.

ber $e^{s \log n}$, where $\log n$ denotes the real value of the logarithm of n. For

$$\frac{(n-1)!n^s}{s(s+1)\cdots(s+n-1)} = \frac{1}{s}\prod_{\nu=1}^{n-1}\frac{\left(1+\frac{1}{\nu}\right)^s}{1+\frac{s}{\nu}},$$

and the infinite product which is the limit, as n approaches ∞, of the right member of this equation, converges uniformly for every s bounded in absolute value and different from $0, -1, -2, \cdots$. An infinite product of the form $\prod_{\nu=1}^{\infty}(1+a_\nu)$ will converge if the series $\sum_{\nu=1}^{\infty} a_\nu$ is absolutely convergent and therefore, in particular, if $\nu^2|a_\nu|$ approaches a finite limit. But, given any number $N > 1$, we have for any s such that $|s| < N$ and any $\nu > N$

$$\left|\frac{\left(1+\frac{1}{\nu}\right)^s}{1+\frac{s}{\nu}} - 1\right| = \frac{\left|\frac{s(s-1)}{2!}\cdot\frac{1}{\nu^2} + \frac{s(s-1)(s-2)}{3!}\cdot\frac{1}{\nu^3} + \cdots\right|}{\left|1+\frac{s}{\nu}\right|}$$

$$< \left(\frac{N(N+1)}{2!}\cdot\frac{1}{\nu^2} + \cdots\right)\Big/\left(1-\frac{N}{\nu}\right),$$

and the product of this last expression (which is independent of s) by ν^2 approaches the limit $\frac{1}{2}N(N+1)$ as ν increases indefinitely. Therefore the infinite product under consideration *converges*, and that *uniformly*, for every s bounded in absolute value, and this remains true when s is zero or a negative integer provided the corresponding infinite factor is omitted from the product.

The function defined by this infinite product is denoted by $\Gamma(s)$, so that

$$\Gamma(s) = \frac{1}{s}\prod_{\nu=1}^{\infty}\frac{\left(1+\frac{1}{\nu}\right)^s}{1+\frac{s}{\nu}} \tag{3}$$

or

$$\Gamma(s) = \lim_{n\to\infty}\frac{(n-1)!n^s}{s(s+1)\cdots(s+n-1)}. \tag{3'}$$

It has the properties[6]

[6] Weierstrass, K., Ueber die Theorie der analytischen Facultäten. Journ. für Math., vol. 51 (1856), pp. 1–60. Reprinted in Abhandlungen aus der Functionenlehre (Berlin, 1886), pp. 181–262 [and *Mathematische Werke*, vol. 1 (Berlin, 1894), pp. 153–211].

(4a) $$\Gamma(s+1) = s\Gamma(s),$$

(4b) $$\lim_{n\to\infty} \frac{\Gamma(s+n)}{(n-1)!n^s} = 1,$$

from which it also follows that

$$\Gamma(1) = 1 \quad \text{and} \quad \Gamma(n) = (n-1)!$$

For, by (3′),

$$\frac{\Gamma(s+1)}{\Gamma(s)} = \lim_{n\to\infty} \left[\frac{(n-1)!n^{s+1}}{(s+1)(s+2)\cdots(s+n)} \div \frac{(n-1)!n^s}{s(s+1)\cdots(s+n-1)} \right] = s,$$

and therefore $\Gamma(s+1) = s\Gamma(s)$. But, this having been established, we have, as in (2) and (b),

(d) $$\Gamma(s) = \frac{\Gamma(s+n)}{(n-1)!n^s} \cdot \frac{(n-1)!n^s}{s(s+1)\cdots(s+n-1)},$$

and therefore, since the limit of the second factor on the right as n approaches ∞ is itself $\Gamma(s)$, we have, for *all* values of s,

$$\lim_{n\to\infty} \frac{\Gamma(s+n)}{(n-1)!n^s} = 1.$$

That $\Gamma(1) = 1$ is seen at once from (3). It also follows from (4a) and (4b), since when s is a positive integer we may, by (4a), replace $\Gamma(s+n)$ by $(s+n-1)!\Gamma(1)$ in (4b), and (4b) then becomes $\Gamma(1) = 1$.

Conversely the two conditions (4a), (4b) serve to define $\Gamma(s)$. For (d) follows from (4a) and from (d) in turn it follows, by (4b), that $\Gamma(s)$ is the function defined by (3′).

Thus $\Gamma(s)$ is itself a particular solution of the initial functional equation, $f(s+1) = sf(s)$, (1). Notice also that if in (c) we replace $F(s)$ by $\Gamma(s)$ and therefore $F(1)$ by 1, (c) reduces to (3′).

Multiplying[5] (3′) by

$$1 = \lim_{n\to\infty} (n+a)^s/n^s,$$

where a is independent of n but may depend on s, it is seen that

(3″) $$\Gamma(s) = \lim_{n\to\infty} \frac{(n-1)!(n+a)^s}{s(s+1)\cdots(s+n-1)},$$

which is slightly more general in form than (3′) and will be used later.

Before proceeding further, we shall transform (3). It is well known that $1/(\nu+1) < \log(1+1/\nu) < 1/\nu$, or

$$0 < \frac{1}{\nu} - \log\left(1 + \frac{1}{\nu}\right) < \frac{1}{\nu} - \frac{1}{\nu+1},$$

whence it follows that $\Sigma(1/\nu - \log(1 + 1/\nu))$ is convergent. On the other hand,

$$\sum_1^n \left(\frac{1}{\nu} - \log\left(1 + \frac{1}{\nu}\right)\right) = 1 + \tfrac{1}{2} + \cdots + \frac{1}{n} - \log(n+1),$$

and hence we may write

$$C = \sum_1^\infty \left(\frac{1}{\nu} - \log\left(1 + \frac{1}{\nu}\right)\right)$$
$$= \lim_{n\to\infty}\left[1 + \tfrac{1}{2} + \cdots + \frac{1}{n} - \log(n+1)\right],$$

which is known as Euler's constant.[7] Its value correct to ten decimals is $0.5772156649\cdots$.[8] Consequently

$$e^{-Cs} = e^{\sum_1^\infty \left(s\log\left(1+\frac{1}{\nu}\right) - \frac{s}{\nu}\right)} = \prod_1^\infty \left(1 + \frac{1}{\nu}\right)^s e^{-s/\nu},$$

and dividing this by (3), we obtain

$$\frac{1}{\Gamma(s)} = e^{Cs}\cdot s\prod_1^\infty\left(1 + \frac{s}{\nu}\right)e^{-s/\nu}. \tag{5}$$

As stated above, the condition (4*b*) was given by Weierstrass in his famous memoir on analytic factorials,[6] which may be considered as an unsurpassed model of rigorous and clear exposition. Equation (3) is due to Euler,[9] who unfortunately soon replaced this excellent definition by definite integrals[10]; in consequence, several of the formal properties of the Gamma function escaped his attention. (3′) is due to Gauss[11] who undoubtedly was not familiar with Euler's expression. (5) is due to [Schlömilch and] Newman[12], and was also established by Weierstrass.[6] Euler's constant, referred to above, is sometimes, without any justification, called Mascheroni's constant. What we have denoted here by $\Gamma(s)$ according to

---

[7] Euler, L. De progressionibus harmonicis observationes. Comment. Acad. Petrop. vol. 7 (1734–1735, published 1740), p. 156.

[8] [The arithmetical character of this constant is entirely unknown; it has not even been shown to be irrational.]

[9] Euler, L. Letter to Goldbach, Oct. 13, 1729. Correspondence math. et phys. de quelques célèbres géomètres du 18e siècle, publiée par Fuss, vol. 1 (St. Pétersburg, 1843), p. 1.

[10] Euler, L. De progressionibus transcendentibus seu quarum termini generales algebraice dari nequeunt. Comment. Acad. Petrop., vol. 5 (1730–31, published 1738), pp. 36–57.

[11] Gauss, C. F. Disquisitiones generales circa seriem infinitam etc. Comment. Gotting., vol. 2 (1813), pp. 1–46. Reprinted in Werke, vol. 3 (1876), pp. 122–162. [German translation by H. Simon, Berlin, 1888.]

[12] [Schlömilch, O. Einiges über die Eulerschen Integrale der zweiten Art. Grunert Archiv, vol. 4 (1844), pp. 167–174.] Newman, On $\Gamma(a)$ especially when a is negative. Cambridge and Dublin math. Journal, vol. 3 (1848), pp. 57–60.

Legendre,[13] was written $[s-1]$ by Euler, $\Pi(s-1)$ by Gauss and $1/Fc(s)$ by Weierstrass.

2. Properties. From (5) it follows immediately, by well known elementary propositions, that $1/\Gamma(s)$ *may be expanded in a power series convergent for all finite values of* s, or that $1/\Gamma(s)$ is an entire transcendental function, with the simple zeros $0, -1, -2, \cdots$.[6] The function $\Gamma(s)$ is consequently a single-valued analytic function of s with the poles of order one $0, -1, -2, \cdots$, and it is seen at the same time that the equation $\Gamma(s) = 0$ has no roots. All this also follows from the fact that the product on the right side of (4) was shown to be uniformly convergent.

Since $\sin \pi s$ has the zeros $0, \pm 1, \pm 2, \cdots$, one is led to consider the function $1/(\Gamma(s)\Gamma(1-s))$ which has the same zeros. By reason of (4*a*) and (3) or (5) it is found that

$$\frac{1}{\Gamma(s)\Gamma(1-s)} = -\frac{1}{s\Gamma(s)\Gamma(-s)} = s\prod_{1}^{\infty}\left(1-\frac{s^2}{\nu^2}\right).$$

This product, however, equals $(1/\pi)\sin \pi s$, by a well known theorem due to Euler and proved in the elements of the theory of functions, and consequently

$$\Gamma(s)\Gamma(1-s) = \frac{\pi}{\sin \pi s}, \tag{6}$$

which is also due to Euler.[14] For $s = \frac{1}{2}$, this formula gives

$$\Gamma(\tfrac{1}{2}) = \sqrt{\pi}, \tag{7}$$

where the square root is positive, since the product expansions (3) and (5) both show that *when s is real and positive, the same is true of* $\Gamma(s)$. Equation (7) was discovered by Euler as early as 1729.[9] Since it is known that[15]

---

[13] [Legendre, A. M. Recherches sur diverses sortes d'intégrales définies, Mém. de l'Institut, vol. 10 (1809), pp. 416–509.]

[14] [Euler, L. Evolutio formulæ integralis, etc. Novi Comment. Acad. Petrop. vol. 16 (1771, published 1772), pp. 91–139.]

[15] {For when h is odd,

$$\begin{aligned}\sin\frac{\pi}{h}\sin\frac{2\pi}{h}\cdots\sin\frac{(h-1)\pi}{h} &= \sin^2\frac{\pi}{h}\sin^2\frac{2\pi}{h}\cdots\sin^2\frac{h-1}{2h}\pi \\ &= 2^{(1-h)/2}\left(1-\cos\frac{2\pi}{h}\right)\left(1-\cos\frac{4\pi}{h}\right)\cdots\left(1-\cos\frac{(h-1)\pi}{h}\right).\end{aligned} \tag{a}$$

But h being odd we have

$$x^h + 1 = (x+1)\prod_{i=1}^{(h-1)/2}\left(x^2 - 2x\cos\frac{2i-1}{h}\pi + 1\right)$$

and if both members of this identity be divided by $x+1$ and x be then set equal to 1 in the result we obtain

$$h = 2^{(h-1)/2}\left(1-\cos\frac{2\pi}{h}\right)\left(1-\cos\frac{4\pi}{h}\right)\cdots\left(1-\cos\frac{(h-1)\pi}{h}\right). \tag{b}$$

By substituting (*b*) in (*a*) we at once obtain the formula of the text. The proof for h even is similar.}

$$2^{1-h} \cdot h = \sin\frac{\pi}{h}\sin\frac{2\pi}{h}\cdots\sin\frac{(h-1)\pi}{h},$$

where h is a positive integer, we find by making $s = 1/h, 2/h, \cdots, (h-1)/h$ in (6), multiplying the resulting equations and extracting the square root,

$$\Gamma\left(\frac{1}{h}\right)\Gamma\left(\frac{2}{h}\right)\cdots\Gamma\left(\frac{h-1}{h}\right) = (2\pi)^{(h-1)/2}h^{-1/2}, \tag{8}$$

which is also due to Euler. In § 4, we shall give another proof of this equation without using the properties of the sine. Writing

$$(s, n) = \frac{\Gamma(n)n^s}{s(s+1)\cdots(s+n-1)},$$

it follows from (3′) that $\lim\limits_{n\to\infty}(s, n) = \Gamma(s)$; it is readily seen that

$$\frac{(s, n)\left(s+\frac{1}{h}, n\right)\cdots\left(s+\frac{h-1}{h}, n\right)}{(hs, hn)} = \frac{\Gamma(n)^h n^{(h-1)/2}h^{hn}}{\Gamma(nh)}\cdot h^{-hs},$$

and since the left side of this equation has a definite limit for $n \to \infty$, the same must be true for the right side, so that

$$\frac{\Gamma(s)\Gamma\left(s+\frac{1}{h}\right)\cdots\Gamma\left(s+\frac{h-1}{h}\right)}{\Gamma(hs)} = k\cdot h^{-hs},$$

where the constant k may be determined with the aid of (8) by making $s = 1/h$, which gives $k = (2\pi)^{(h-1)/2}h^{1/2}$. Thus we have proved Gauss's[11] theorem[16, 17]

$$\Gamma(s)\Gamma\left(s+\frac{1}{h}\right)\cdots\Gamma\left(s+\frac{h-1}{h}\right) = h^{1/2-hs}(2\pi)^{(h-1)/2}\Gamma(hs). \tag{9}$$

[16] [The special case $h = 2$ is due to Legendre.[13]]

[17] [This may be generalized as follows: h and k being positive integers, (9) gives

$$\Gamma\left(hs+\frac{h\nu}{k}\right) = h^{h(s+\nu/k)-1/2}(2\pi)^{-(h-1)/2}\prod_{\mu=0}^{h-1}\Gamma\left(s+\frac{\nu}{k}+\frac{\mu}{h}\right),$$

whence

$$\prod_{\nu=0}^{k-1}\Gamma\left(hs+\frac{h\nu}{k}\right) = h^{hks+[h(k-1)]/2-k/2}(2\pi)^{-[k(h-1)]/2}\prod_{\nu=0}^{k-1}\prod_{\mu=0}^{h-1}\Gamma\left(s+\frac{\nu}{k}+\frac{\mu}{h}\right).$$

Dividing this equation by the one obtained by interchanging h and k, μ and ν, it is seen that

$$\frac{\prod_{\nu=0}^{k-1}\Gamma\left(hs+\frac{h\nu}{k}\right)}{\prod_{\mu=0}^{h-1}\Gamma\left(ks+\frac{k\mu}{h}\right)} = \left(\frac{h}{k}\right)^{hks+(hk-h-k)/2}\cdot(2\pi)^{(k-h)/2}, \tag{9′}$$

which reduces to (9) for $k = 1$. Equation (9′) was obtained from the definite integral expression for $\log\Gamma(s)$ by Winckler, A., Neue Theoreme zur Lehre von den bestimmten Integralen, Sitzungsber. Akad. Wien, vol. 21 (1856), pp. 389–426. Compare Nielsen,[3] pp. 196–198 and 326. The remark that (9′) is a consequence of (9) appears to be new.]

3. Evaluation of some infinite products by means of the Gamma function. In the notation introduced above, we have for a positive integer $m > n$,

$$\frac{(s,\, m)}{(s,\, n)} = \frac{\left(\frac{m}{n}\right)^s}{\left(1+\frac{s}{n}\right)\left(1+\frac{s}{n+1}\right)\cdots\left(1+\frac{s}{m-1}\right)},$$

whence, letting m and n increase indefinitely in a quite arbitrary fashion,

$$\frac{\Gamma(s)}{\Gamma(s)} = 1 = \lim \frac{\left(\frac{m}{n}\right)^s}{\left(1+\frac{s}{n}\right)\left(1+\frac{s}{n+1}\right)\cdots\left(1+\frac{s}{m-1}\right)}.$$

Assuming in particular that m/n *tends toward a definite limit*, we have

$$\text{(10)} \qquad \lim \left(\frac{m}{n}\right)^s = \lim \prod_n^{m-1}\left(1+\frac{s}{\nu}\right).$$

By division of Gamma functions of different variables we obtain from either of equations (3) and (5)

$$\text{(11)} \qquad \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+s)\Gamma(\beta-s)} = \prod_0^\infty\left(1+\frac{s}{\alpha+\nu}\right)\left(1-\frac{s}{\beta+\nu}\right).$$

Writing, according to Mellin,[18] $R(s) = a_1 s + a_2 s^2 + \cdots + a_n s^n$, and denoting the n roots of the equation

$$\rho^n(1 + R(1/\rho)) = \rho^n + a_1\rho^{n-1} + \cdots + a_n = 0$$

by $\rho_1, \rho_2, \cdots, \rho_n$, it follows from (5) that

$$\text{(12)} \qquad \begin{aligned} &\frac{\Gamma(\alpha)^n}{\Gamma(\alpha-\rho_1 s)\Gamma(\alpha-\rho_2 s)\cdots\Gamma(\alpha-\rho_n s)} \\ &\qquad = e^{Ca_1 s}\left(1+R\left(\frac{s}{\alpha}\right)\right)\prod_1^\infty\left(1+R\left(\frac{s}{\alpha+\nu}\right)\right)e^{-a_1 s/\nu}. \end{aligned}$$

We note the special case $R(s) = -s^n$, which gives for $\alpha = 1$

$$\begin{aligned} &\frac{1}{\Gamma(1-s)\Gamma(1-e^{2\pi i/n}s)\Gamma(1-e^{4\pi i/n}s)\cdots\Gamma(1-e^{[(n-1)2\pi i]/n}s)} \\ &\qquad = \prod_1^\infty\left(1-\left(\frac{s}{\nu}\right)^n\right), \end{aligned}$$

a formula indicated by Liouville.[19]

[18] Mellin, H., Eine Verallgemeinerung der Gleichung $\Gamma(x)\Gamma(1-x) = \pi/\sin \pi x$. Acta Math., vol. 3 (1883), pp. 102–104. Über gewisse durch die Gammafunktion ausdrückbare Producte, ibid., pp. 322–324.

[19] Liouville, J., Note sur la fonction gamma de Legendre. Comptes rendus vol. 35 (1852); pp. 317–322, and Journal de math., vol. 17 (1852), pp. 448–453.

When $R(s)$ is an *even* function of s, the convergent product

$$\prod_0^\infty\left(1+R\left(\frac{s}{\nu+1}\right)\right)$$

is expressible by trigonometric and exponential functions alone, on account of (6). This result is also easily derived from Euler's product expansion of $\sin \pi x$ without introducing the Gamma function.

4. Expansion of $\log \Gamma(s)$. Any finite and definite limit $\lim\limits_{n\to\infty} u_n$ may be rewritten as an infinite *convergent* series $u_h + \sum\limits_h^\infty (u_{\nu+1} - u_\nu)$. This remark will now be used to transform equation (4″), by taking logarithms on both sides and making $a = s$. It should be noted that in the following any logarithm is defined as that branch of the logarithmic function which is real for real and positive values of the variable. When s is represented in the complex plane in the familiar way, both sides of the equation referred to will then be single-valued as long as we do not make a circuit of any of the points $0, -1, -2, \cdots$. Thus, writing

$$u_n = \log \Gamma(n) + s \log (n+s) - \log s - \log (s+1) - \cdots - \log (s+n-1),$$

we obtain

$$u_{\nu+1} - u_\nu = \log \nu + s \log \frac{s+\nu+1}{s+\nu} - \log (s+\nu)$$

$$= \left[(s+\nu+\tfrac{1}{2}) \log \frac{s+\nu+1}{s+\nu} - 1\right] - \left[(\nu+\tfrac{1}{2}) \log \frac{\nu+1}{\nu} - 1\right]$$

$$- \left[(\nu+\tfrac{1}{2}) \log \frac{s+\nu+1}{\nu+1} - (\nu-\tfrac{1}{2}) \log \frac{s+\nu}{\nu}\right].$$

Now the expression inside the last of the three square brackets is the general term of a *convergent* series, since

$$\lim_{n\to\infty} (n+\tfrac{1}{2}) \log \frac{s+n+1}{n+1} = \lim_{n\to\infty} (n+\tfrac{1}{2}) \log \left(1+\frac{s}{n+1}\right) = s,$$

and the sum of this series for $\nu = 1, 2, 3, \cdots$, therefore, in consequence of the remark made above, equals $s - \frac{1}{2} \log (1+s)$. Furthermore, $\Sigma\left[(\nu+\tfrac{1}{2}) \log \frac{\nu+1}{\nu} - 1\right]$ converges, since for $\nu > 1$,

$$(\nu+\tfrac{1}{2}) \log \left(1+\frac{1}{\nu}\right) - 1 = \frac{1}{12\nu^2} - \frac{1}{12\nu^3} + \cdots,$$

and the product of this expression by ν^2 is bounded. Consequently $\Sigma\left[(s+\nu+\frac{1}{2})\log\frac{s+\nu+1}{s+\nu}-1\right]$ must be convergent, since $\Sigma(u_{\nu+1}-u_\nu)$ converges, and we may write

$$\omega(s)=\sum_0^\infty\left[(s+\nu+\tfrac{1}{2})\log\frac{s+\nu+1}{s+\nu}-1\right]. \tag{13}$$

It is then seen at once that

$$\begin{aligned}\text{og}\,\Gamma(s) &= u_1+\sum_1^\infty(u_{\nu+1}-u_\nu)\\ &= [s\log(1+s)-\log s]+\left[\omega(s)-(s+\tfrac{1}{2})\log\frac{s+1}{s}+1\right]\\ &\qquad -\omega(1)-[s-\tfrac{1}{2}\log(1+s)]\\ &= (s-\tfrac{1}{2})\log s-s+1-\omega(1)+\omega(s).\end{aligned}$$

It remains to determine the constant $1-\omega(1)$, which we denote by k' for brevity, and to this purpose we observe that $\lim\limits_{n\to\infty}\omega(s+n)=0$, since $\omega(s+n)$ is *the remainder in the convergent series* $\omega(s)$, counted from the nth term. Making $s=n$ and $s=2n$ respectively, it is found that

$$\begin{aligned}2\log\Gamma(n) &= (2n-1)\log n-2n+2k'+2\omega(n),\\ \log\ \Gamma(2n) &= (2n-\tfrac{1}{2})\log 2n-2n+k'+\omega(2n),\end{aligned}$$

whence by subtraction

$$\begin{aligned}k'+2\omega(n)-\omega(2n) &= \log\frac{\Gamma(n)^2n^{1/2}2^{2n-1/2}}{\Gamma(2n)}\\ &= \log\frac{2^{1/2}\Gamma(n)n^{1/2}}{\frac{1}{2}(\frac{1}{2}+1)\cdots(\frac{1}{2}+n-1)},\end{aligned}$$

and increasing n indefinitely

$$k'=\log[2^{1/2}\Gamma(\tfrac{1}{2})]=\log\sqrt{2\pi}.$$

Hence the final result becomes

$$\log\Gamma(s)=(s-\tfrac{1}{2})\log s-s+\log\sqrt{2\pi}+\omega(s). \tag{14}$$

The infinite series $\omega(s)$ is due to Gudermann.[20] As special cases of (14) we note, for $s=1$ and $s=\frac{1}{2}$ respectively,

$$\omega(1)=1-\log\sqrt{2\pi}\qquad\text{and}\qquad\omega(\tfrac{1}{2})=\tfrac{1}{2}(1-\log 2).$$

[20] Gudermann, C., Additamentum ad functionis $\Gamma(a)$ theoriam. Journal für Math., vol. 29 (1845), pp. 209–212.

We have seen above that $\omega(s)$ converges toward zero when s increases by positive integral increments toward infinity. Therefore, by (14), $(s - \frac{1}{2}) \log s - s + \log \sqrt{2\pi}$ becomes an *approximate expression* for $\log \Gamma(s)$ when s increases in the manner indicated. When s is a positive integer, we have, in particular, the approximation formula known as Stirling's[21] formula. Before proceeding to generalize this result, we shall give some applications of (14). We have

$$\Gamma(s) = \lim_{n \to \infty} \frac{\Gamma(n) n^s}{s(s+1) \cdots (s+n-1)}$$

$$= \lim_{n \to \infty} \frac{n^{n-1/2} \cdot e^{-r} \cdot \sqrt{2\pi} \cdot n^s \cdot e^{\omega(n)}}{s(s+1) \cdots (s+n-1)},$$

and consequently

$$\Gamma(s) = \sqrt{2\pi} \lim_{n \to \infty} \frac{n^{n+s-1/2} e^{-n}}{s(s+1) \cdots (s+n-1)}, \tag{15}$$

which is due to [Enneper and] Gilbert.[22] Furthermore we may determine by means of (14), as shown by Cauchy,[23] the limit k of

$$\frac{\Gamma(n)^h n^{h-1/2} h^{hn}}{\Gamma(nh)} = h^{1/2} (2\pi)^{h-1/2} e^{h\omega(n) - \omega(nh)},$$

or $k = h^{1/2}(2\pi)^{h-1/2}$. Thus Gauss's theorem (9), and (8) as a particular case of it, is proved anew. The most elegant proof is however obtained from (15). The notation $\omega(s)$ is due to Catalan; Binet[24] denoted this function by $\mu(s)$ and Cauchy by $\varpi(s)$.

5. Investigation of $\omega(s)$ and $\log \Gamma(s)$ for large values of $|s|$. When $|\alpha + 1| > 1$, we have

$$(\alpha + \tfrac{1}{2}) \log\left(1 + \frac{1}{\alpha}\right) - 1 = -[(\alpha + 1) - \tfrac{1}{2}] \log\left(1 - \frac{1}{\alpha + 1}\right) - 1$$

$$= \sum_{\nu=2}^{\infty} \frac{\nu - 1}{2\nu(\nu + 1)} \frac{1}{(\alpha + 1)^\nu}.$$

---

[21] Stirling, J. Methodus differentialis sive tractatus de summatione et interpolatione serierum infinitarum, London, 1730, p. 135.

[22] [Enneper, A., Über die Function Π von Gauss mit complexem Argument. Diss. Göttingen 1856.] Gilbert, Ph., Recherches sur le développement de la fonction Γ et sur certaines intégrales définies qui en dépendent. Mém. Ac. Belgique, vol. 41 (1873), pp. 1–60.

[23] [Cauchy, A. L., Exercices de math., vol. 2 (Paris, 1827), pp. 91–92, and Nouveaux exercices, vol. 2 (Paris, 1841), pp. 407–408.]

[24] Binet, J., Mémoire sur les intégrales Eulériennes et sur leur application à la théorie des suites, ainsi qu'à l'évaluation des fonctions de grands nombres. Journal de l'École Polytechnique, cahier 27 (1839), pp. 123–343.

Now $\dfrac{\nu - 1}{2\nu(\nu + 1)} \leqq \frac{1}{12}$ for $\nu = 2, 3, 4, \cdots$, and consequently

$$\left|(\alpha + \tfrac{1}{2}) \log\left(1 + \frac{1}{\alpha}\right) - 1\right| < \frac{1}{12}\sum_{\nu=2}^{\infty} \frac{1}{|\,\alpha + 1\,|^{\nu}} = \frac{1}{12}\, \frac{1}{|\,\alpha + 1\,|(|\,\alpha + 1\,| - 1)}.$$

Applying this to the series (13), and observing that

$$\Sigma \frac{1}{|\,s + \nu + 1\,|\,(\,|\,s + \nu + 1\,| - 1)}$$

must be convergent since $\nu^2/[|\,s + \nu + 1\,|\,(|\,s + \nu + 1\,| - 1)]$ is bounded for ν increasing, it follows first that $\omega(s)$ may be expanded in the *convergent double series* (Binet[24])

$$\begin{aligned} \omega(s) &= \sum_{\nu=2}^{\infty} \frac{\nu - 1}{2\nu(\nu + 1)} \sum_{\mu=1}^{\infty} \frac{1}{(s + \mu)^{\nu}} \\ &= \frac{1}{12}\sum_{\mu=1}^{\infty} \frac{1}{(s + \mu)^2} + \frac{1}{12}\sum_{\mu=1}^{\infty} \frac{1}{(s + \mu)^3} + \cdots, \end{aligned} \tag{16}$$

and second, that

$$|\,\omega(s)\,| < \frac{1}{12}\sum_{\nu=1}^{\infty} \frac{1}{|\,s + \nu\,|\,(\,|\,s + \nu\,| - 1)},$$

all this under the assumption that $|\,s + 1\,|$, $|\,s + 2\,|$, $\cdots$, are all greater than unity, or that s, when represented as usual in the complex plane, lies outside a sequence of circles with the radius 1 and the points -1, -2, -3, as centers, which we shall assume here for the sake of simplicity. (If s is inside or on the circumference of one or two of these circles, it is only necessary to omit one or two corresponding terms in the series for $\omega(s)$.)

Before proceeding to consider the last inequality for complex values of s, we shall make an observation connected with Stirling's formula. When s is real and positive, we have

$$0 < \omega(s) < \frac{1}{12}\sum_{1}^{\infty} \frac{1}{(s + \nu - 1)(s + \nu)} = \frac{1}{12}\sum_{1}^{\infty}\left(\frac{1}{s + \nu - 1} - \frac{1}{s + \nu}\right) = \frac{1}{12s},$$

and consequently by (14)

$$\log \Gamma(s) = (s - \tfrac{1}{2}) \log s - s + \log \sqrt{2\pi} + \frac{\theta}{12s}, \qquad 0 < \theta < 1,$$

a proposition which is frequently used and is commonly proved by the integral calculus.[25]

[25] [A more accurate formula is readily obtained by evaluating a lower bound for $\omega(s)$ in the same manner as the upper bound was found above. From (16) we obtain, s being real and positive,

$$\omega(s) > \frac{1}{12}\sum_{1}^{\infty} \frac{1}{(s + \nu)^2} > \frac{1}{12}\sum_{1}^{\infty} \frac{1}{(s + \nu)(s + \nu + 1)} = \frac{1}{12}\cdot\frac{1}{s + 1},$$

In the general case, let $(1+\rho)$ be the shortest distance from $s = x + yi$ to any of the points $-1, -2, -3, \cdots$; then ρ is positive according to our hypothesis and $|s+\nu| > 1+\rho$, so that

$$\frac{1}{|s+\nu|(|s+\nu|-1)} \leqq \frac{1}{|s+\nu|^2} \cdot \frac{1}{1-\dfrac{1}{1+\rho}} = \frac{1+\rho}{\rho} \cdot \frac{1}{|s+\nu|^2}$$
$$= \frac{1+\rho}{\rho} \frac{1}{x^2+y^2+2\nu x+\nu^2}.$$

First supposing x *positive*, and noting that $A^2 + B^2 \geqq \frac{1}{2}(A+B)^2$ when A and B are real and positive, the last expression will be less than

$$\frac{1+\rho}{\rho} \frac{1}{|s|^2+\nu^2} \leqq \frac{1+\rho}{\rho} \frac{2}{(|s|+\nu)^2} < \frac{1+\rho}{\rho} \frac{2}{(|s|+\nu-1)(|s|+\nu)}$$
$$= 2\frac{1+\rho}{\rho}\left(\frac{1}{|s|+\nu-1} - \frac{1}{|s|+\nu}\right),$$

and consequently

$$\text{(I)} \qquad |\omega(s)| < \frac{1+\rho}{6\rho} \sum_{\nu=1}^{\infty}\left(\frac{1}{|s|+\nu-1} - \frac{1}{|s|+\nu}\right)$$
$$= \frac{1+\rho}{6\rho} \cdot \frac{1}{|s|}, \qquad (x > 0).$$

On the other hand, let $-(m+1) < x \leqq -m$, where m is a positive integer or zero; then

$$\sum_{\nu=1}^{\infty} \frac{1}{|s+\nu|^2} = \sum_{\nu=1}^{m-1} \frac{1}{|s+\nu|^2} + \sum_{\nu=m}^{m+1} \frac{1}{|s+\nu|^2} + \sum_{\nu=m+2}^{\infty} \frac{1}{|s+\nu|^2},$$

where, for $m = 0$ or 1, the first sum on the right side should be omitted. In the second sum on the right side we have

---

and consequently

$$\log \Gamma(s) = \left(s - \frac{1}{2}\right) \log s - s + \log \sqrt{2\pi} + \frac{1}{12} \cdot \frac{1}{s+\theta}, \qquad 0 < \theta < 1.$$

Retaining two terms in (16), we may even show that $0 < \theta < 1/2$ in the last formula, since

$$\omega(s) > \frac{1}{12} \sum_1^{\infty} \frac{1}{(s+\nu)^2} + \frac{1}{12} \sum_1^{\infty} \frac{1}{(s+\nu)^3}$$

and, for x real and > 1,

$$\frac{1}{x^2} + \frac{1}{x^3} - \frac{1}{x-\frac{1}{2}} + \frac{1}{x+\frac{1}{2}} = \frac{4x^2-x-1}{x^3(4x^2-1)} > 0,$$

whence, making $x = s + \nu$,

$$\omega(s) > \frac{1}{12} \sum_1^{\infty} \left(\frac{1}{s+\nu-\frac{1}{2}} - \frac{1}{s+\nu+\frac{1}{2}}\right) = \frac{1}{12} \cdot \frac{1}{s+\frac{1}{2}}.$$

This proof is believed to be new; for a still closer limitation of θ, obtained by definite integrals, see Sonin, M. N., Sur les termes complémentaires de la formule sommatoire d'Euler et de celle de Stirling, Comptes rendus, vol. 108 (1889), pp. 725–727, and Annales de l'École Normale, ser. 3, vol. 6 (1889), pp. 257–262.]

$$\sum_{\nu=m}^{m+1} \frac{1}{|s+\nu|^2} = \frac{1}{(x+m)^2+y^2} + \frac{1}{(x+m+1)^2+y^2} < \frac{1}{y^2} + \frac{1}{y^2};$$

making $\nu = m - \mu$ in the first sum, so that $\mu = m-1, m-2, \cdots, 2, 1$, and $\nu = m + 1 + \mu$ in the third sum, so that $\mu = 1, 2, 3, \cdots$, we find, since $x + m \leqq 0$, $x + m + 1 > 0$,

$$\sum_{\nu=1}^{m-1} \frac{1}{|s+\nu|^2} = \sum_{\mu=1}^{m-1} \frac{1}{(\mu-(x+m))^2+y^2} \leqq \sum_{\mu=1}^{m-1} \frac{1}{\mu^2+y^2}$$

$$< \sum_{\mu=1}^{\infty} \frac{1}{\mu^2+y^2} \leqq \sum_{\mu=1}^{\infty} \frac{2}{(\mu+|y|)^2}$$

$$< 2\sum_{\mu=1}^{\infty} \frac{1}{(|y|+\mu-1)(|y|+\mu)} = \frac{2}{|y|};$$

$$\sum_{\nu=m+2}^{\infty} \frac{1}{|s+\nu|^2} = \sum_{\mu=1}^{\infty} \frac{1}{(\mu+(x+m+1))^2+y^2}$$

$$< \sum_{\mu=1}^{\infty} \frac{1}{\mu^2+y^2} < \frac{2}{|y|},$$

so that finally, adding the three sums,

$$\text{(II)} \qquad |\omega(s)| < \frac{1+\rho}{6\rho}\left(\frac{2}{|y|} + \frac{1}{y^2}\right), \qquad (x \leqq 0).\text{[26]}$$

Since $(1+\rho)/\rho$ is bounded, it is now seen immediately from (I) and (II) that $\omega(s)$ *converges uniformly toward zero, when the distance between s and the nearest point on the negative real axis increases indefinitely.* In particular, this condition is satisfied when $|s|$ increases indefinitely while the arc of s[27] is constant and different from π. When $|s|$ increases in the manner indicated, it therefore follows from (14) that

$$\text{(17)} \qquad \lim_{|s|\to\infty} \frac{\Gamma(s)}{s^{s-1/2}e^{-s}\sqrt{2\pi}} = 1,$$

where that branch of the multiple-valued function $s^{s-1/2}$ is taken which is real for real and positive values of s. This equation may be used to extend the domain of validity of (4*b*); it is readily seen that, when t is any complex quantity,

$$\lim_{|s|\to\infty} \frac{\Gamma(s+t)}{\Gamma(s)s^t} = 1.$$

We have used above the expansion of $(\alpha + \frac{1}{2}) \log (1 + 1/\alpha) - 1$ in

[26] [The form and proof of (II) have been slightly changed from the original.]

[27] [When s is written in the form $s = |s| e^{\theta i}$, θ is the arc of s].

powers of $1/(\alpha + 1)$; the expansion in powers of $1/\alpha$ would have served the same purpose. In fact, for $|\alpha| > 1$,

$$(\alpha + \tfrac{1}{2}) \log\left(1 + \frac{1}{\alpha}\right) - 1 = \sum_{\nu=2}^{\infty} \frac{(-1)^{\nu}(\nu - 1)}{2\nu(\nu + 1)} \frac{1}{\alpha^{\nu}},$$

whence we deduce, instead of (16), the expansion

$$\begin{aligned} \omega(s) &= \sum_{\nu=2}^{\infty} \frac{(-1)^{\nu}(\nu - 1)}{2\nu(\nu + 1)} \sum_{\mu=0}^{\infty} \frac{1}{(s + \mu)^{\nu}} \\ &= \frac{1}{12} \sum_{\mu=0}^{\infty} \frac{1}{(s + \mu)^2} - \frac{1}{12} \sum_{\mu=0}^{\infty} \frac{1}{(s + \mu)^3} + \cdots, \end{aligned} \tag{18}$$

valid for $|s|, |s + 1|, |s + 2|, \cdots$ all greater than unity. This formula is also due to Binet.[28] Neither (16) nor (18) is adapted to the numerical calculation of $\omega(s)$ (and $\log \Gamma(s)$ through (14)); expansions suitable for this purpose will be derived in § 14.

6. The function $\psi(s)$. Definition and properties. It follows from (3) that

$$\log \frac{\Gamma(s + t)}{\Gamma(s)} = -\log\left(1 + \frac{t}{s}\right) + \sum_{1}^{\infty}\left[t \log\left(1 + \frac{1}{\nu}\right) - \log\left(1 + \frac{t}{s + \nu}\right)\right],$$

which series is uniformly convergent, since it is derived from uniformly convergent infinite products. When s is not equal to zero or a negative integer, $\Gamma(s + t)$ may be expanded in a Taylor series $\Gamma(s + t) = \Gamma(s) + t\Gamma'(s) + \cdots$ for sufficiently small values of $|t|$. Substituting this on the left and expanding both sides in powers of t, we find by comparison of the coefficients of t

$$\frac{\Gamma'(s)}{\Gamma(s)} = -\frac{1}{s} + \sum_{1}^{\infty}\left[\log\left(1 + \frac{1}{\nu}\right) - \frac{1}{s + \nu}\right].$$

This function, which will be denoted by $\psi(s)$, may be expressed in the following forms, using the series

$$C = \sum_{1}^{\infty}\left[\frac{1}{\nu} - \log\left(1 + \frac{1}{\nu}\right)\right]$$

and

$$-\frac{1}{s} = \sum_{0}^{\infty}\left(\frac{1}{s + \nu + 1} - \frac{1}{s + \nu}\right),$$

$$\begin{aligned} \psi(s) &= -C - \frac{1}{s} + \sum_{1}^{\infty}\left(\frac{1}{\nu} - \frac{1}{s + \nu}\right) \\ &= -C + \sum_{0}^{\infty}\left(\frac{1}{\nu + 1} - \frac{1}{s + \nu}\right). \end{aligned} \tag{19}$$

[28] Binet, J., Abstract of paper quoted in [24], Comptes rendus, vol. 9 (1839), pp. 39–45.

These series are uniformly convergent for all finite values of s; for choosing an N so large that $N > |s|$, it is seen that for $\nu > N$

$$\left|\frac{1}{\nu+1} - \frac{1}{s+\nu}\right| = \left|\frac{s-1}{(\nu+1)(s+\nu)}\right| < \frac{N+1}{(\nu+1)(\nu-N)},$$

the last expression being independent of s and being the general term of a convergent series. Since

$$-\frac{1}{s} + \sum_{1}^{n-1}\left[\log\left(1+\frac{1}{\nu}\right) - \frac{1}{s+\nu}\right] = \log n - \frac{1}{s} - \frac{1}{s+1} - \cdots - \frac{1}{s+n-1},$$

we have

$$(19') \qquad \psi(s) = \lim_{n\to\infty}\left(\log n - \frac{1}{s} - \frac{1}{s+1} - \cdots - \frac{1}{s+n-1}\right),$$

and using $0 = \lim\limits_{n\to\infty} [\log(n+a) - \log n]$, this may also be written

$$(19'') \quad \psi(s) = \lim_{n\to\infty}\left(\log(n+a) - \frac{1}{s} - \frac{1}{s+1} - \cdots - \frac{1}{s+n-1}\right).$$

From (19) we furthermore obtain

$$(19''') \qquad \psi(s) - \psi(t) = \sum_{0}^{\infty}\left(\frac{1}{t+\nu} - \frac{1}{s+\nu}\right) = \sum_{0}^{\infty}\frac{s-t}{(t+\nu)(s+\nu)},$$

which is somewhat more general in form than (19).

From either of equations (19) it is seen at once that

$$\psi(1) = -C,$$

and that

$$(20) \qquad \begin{aligned} \psi(s+1) - \psi(s) &= \frac{1}{s}, \\ \psi(s+n) - \psi(s) &= \frac{1}{s} + \frac{1}{s+1} + \cdots + \frac{1}{s+n-1}. \end{aligned}$$

Substituting the last expression in (19′), we find

$$(21) \qquad \lim_{n\to\infty} [\psi(s+n) - \log n] = 0,$$

which will be generalized below. Conversely, it is obvious that (19′) may be derived from (20) and (21), so that *these two equations are sufficient to define* $\psi(s)$.

The expansions of $\psi(s)$ being uniformly convergent, $\psi(s)$ is *a single-valued analytic function of s which has poles of the first order at the points* $s = 0, -1, -2, \cdots$, *but is holomorphic for all other finite values of s.* By (19‴) we have, for $t = 1 - s$,

$$\psi(s) - \psi(1-s) = -\sum_0^\infty\left(\frac{1}{s+\nu} - \frac{1}{\nu+1-s}\right)$$

$$= -\frac{1}{s} + \frac{1}{1-s} - \frac{1}{1+s} + \frac{1}{2-s} - \frac{1}{2+s} + \cdots,$$

and consequently, by a well-known theorem due to Euler,*

(22) $$\psi(s) - \psi(1-s) = -\pi\cot\pi s,$$

which could also be derived from (6) in the same way as (19) was deduced from (3). Writing, for brevity,

$$[s, n] = \log n - \frac{1}{s} - \frac{1}{s+1} - \cdots - \frac{1}{s+n-1},$$

it is readily seen that, if h be any positive integer,

$$h[hs, hn] - [s, n] - \left[s + \frac{1}{h}, n\right] - \cdots - \left[s + \frac{h-1}{h}, n\right] = h\log h,$$

whence, letting n increase indefinitely,

(23) $$h\psi(hs) - \psi(s) - \psi\left(s + \frac{1}{h}\right) - \cdots - \psi\left(s + \frac{h-1}{h}\right) = h\log h.$$

These propositions could all be obtained by logarithmic differentiation of the corresponding ones for $\Gamma(s)$ (Gauss[6]). I have preferred the above elementary deduction, which is as simple as could be desired. What we have denoted here by $\psi(s)$, according to Cauchy, was written $\Psi(s-1)$ by Gauss,[11] $\varphi(s)$ and $z'(s)$ by Legendre[13] and $\lambda'(s)$ by Binet.[24]

The series for $\psi(s)$ being uniformly convergent, $\psi(s+t)$ may be expanded in ascending powers of t for $|t| < |s|, |s+1|, |s+2|, \cdots,$ and this may be effected by expanding the individual terms on the right side of (19). By Taylor's formula, the comparison of the coefficients of t^m shows that

(24) $$\psi^{(m)}(s) = (-1)^{m+1}m!\sum_{\nu=0}^\infty\frac{1}{(s+\nu)^{m+1}}, \qquad (m \geqq 1).$$

These functions are readily seen to be completely defined by the equations

$$\psi^{(m)}(s+1) - \psi^{(m)}(s) = \frac{(-1)^m m!}{s^{m+1}}, \qquad \text{and} \qquad \lim_{n\to\infty}\psi^{(m)}(s+n) = 0.$$

* {For

$$\sin\pi s = \pi s\prod_{\nu=1}^\infty\left(1 - \frac{s^2}{\nu^2}\right).$$

Hence

$$-\pi\cot\pi s = -\frac{d}{ds}\log\sin\pi s = -\frac{1}{s} + \sum_{\nu=1}^\infty\left(\frac{1}{\nu-s} - \frac{1}{\nu+s}\right).\}$$

They furthermore satisfy relations analogous to (22) and (23), which we shall not, however, stop to prove.

7. Evaluation of some infinite series and products by means of $\psi(s)$. By the use of the functions $\psi(s)$, $\psi'(s)$, $\cdots$ which we have introduced, it is easy to find the sum of series of very general form. For instance, when $F(s)$ is a rational function of the form

$$F(s) = \frac{as^{n-1} + bs^{n-2} + \cdots + l}{s^n + \alpha s^{n-1} + \beta s^{n-2} + \cdots + \lambda},$$

the sum of the convergent series $\Sigma[F(s+\nu) - F(t+\nu)]$ is found by resolving $F(s+\nu)$ and $F(t+\nu)$ into a sum of partial fractions in the familiar way, thus decomposing the series into several others which may be evaluated individually by means of (19‴) and (24). When $a = 0$, the series $\Sigma F(s+\nu)$ becomes convergent. In the particular case when $F(s)$ is a rational function of s^2, it is seen incidentally that $\Sigma F(s+\nu)$ may be evaluated without the aid of $\psi(s)$, $\psi'(s)$, $\cdots$ by the exclusive use of Euler's partial fraction expansions for the trigonometric functions. As an example of the application of these remarks we may mention

$$\sum_0^\infty \frac{1}{(s+\nu-\rho_1)(s+\nu-\rho_2)\cdots(s+\nu-\rho_n)}$$
$$= -\frac{\psi(s-\rho_1)}{(\rho_1-\rho_2)(\rho_1-\rho_3)\cdots(\rho_1-\rho_n)} - \frac{\psi(s-\rho_2)}{(\rho_2-\rho_1)(\rho_2-\rho_3)\cdots(\rho_2-\rho_n)} - \cdots - \frac{\psi(s-\rho_n)}{(\rho_n-\rho_1)(\rho_n-\rho_2)\cdots(\rho_n-\rho_{n-1})},$$

where all of $\rho_1, \rho_2, \cdots, \rho_n$ are *distinct*.*

* {Let

$$f(s) = \prod_{i=1}^{n} (s-\rho_i).$$

By a theorem of Euler,

$$\sum_{i=1}^{n} 1/f'(\rho_i) = 0.$$

Hence

$$-\sum_{\nu=0}^{\infty} \frac{1}{(s+\nu-\rho_1)\cdots(s+\nu-\rho_n)} = -\sum_{\nu=0}^{\infty}\sum_{i=0}^{n} \frac{1}{f'(\rho_i)(s+\nu-\rho_i)}$$
$$= \sum_{\nu=0}^{\infty}\sum_{i=1}^{n}\left[\frac{1}{\nu+1} - \frac{1}{s+\nu-\rho_i}\right]\frac{1}{f'(\rho_i)} = \sum_{i=1}^{n}\sum_{\nu=0}^{\infty}\left[\frac{1}{\nu+1} - \frac{1}{s+\nu-\rho_i}\right]\frac{1}{f'(\rho_i)}$$
$$= \sum_{i=1}^{n} [\psi(s-\rho_i) + C]\frac{1}{f'(\rho_i)} = \sum_{i=1}^{n} \frac{\psi(s-\rho_i)}{f'(\rho_i)}.\}$$

It is seen from (5) that

$$\frac{\Gamma(s)}{\Gamma(s+t)} = e^{Ct}\frac{s+t}{s}\prod_1^\infty \frac{1+\dfrac{s+t}{\nu}}{1+\dfrac{s}{\nu}}e^{-t/\nu} = e^{Ct}\left(1+\frac{t}{s}\right)\prod_1^\infty\left(1+\frac{t}{s+\nu}\right)e^{-t/\nu}.$$

Furthermore, it follows from (19) that

$$e^{t\psi(s)} = e^{-Ct-t/s}\prod_1^\infty e^{t/\nu-[t/(s+\nu)]},$$

and consequently (Mellin[29])

$$e^{t\psi(s)}\frac{\Gamma(s)}{\Gamma(s+t)} = \prod_0^\infty\left(1+\frac{t}{s+\nu}\right)e^{-t/(s+\nu)}, \tag{25}$$

a generalization of (5) and reducing to the latter for $s = 1$. Making $t = 1$, we obtain

$$e^{\psi(s)} = s\prod_0^\infty\left(1+\frac{1}{s+\nu}\right)e^{-1/(s+\nu)}. \tag{26}$$

8. Investigation of $\psi(s)$ for large values of $|s|$. Writing

$$u_n = \log(n+s) - \frac{1}{s} - \frac{1}{s+1} - \cdots - \frac{1}{s+n-1}, \qquad u_0 = \log s,$$

it follows from (19″) that for $a = s$

$$\psi(s) = \lim_{n\to\infty} u_n = u_0 + \sum_0^\infty (u_{\nu+1} - u_\nu),$$

and since

$$u_{\nu+1} - u_\nu = \log\left(1+\frac{1}{s+\nu}\right) - \frac{1}{s+\nu},$$

it is seen that, making

$$\omega^*(s) = \sum_0^\infty\left(\frac{1}{s+\nu} - \log\left(1+\frac{1}{s+\nu}\right)\right), \tag{27}$$

we have

$$\psi(s) = \log s - \omega^*(s). \tag{28}$$

When $|\alpha+1| > 1$, then

$$\frac{1}{\alpha} - \log\left(1+\frac{1}{\alpha}\right) = \frac{1}{(\alpha+1)-1} + \log\left(1-\frac{1}{\alpha+1}\right) = \sum_2^\infty\left(1-\frac{1}{\nu}\right)\frac{1}{(\alpha+1)^\nu},$$

and consequently

$$\left|\frac{1}{\alpha} - \log\left(1+\frac{1}{\alpha}\right)\right| < \sum_2^\infty \frac{1}{|\alpha+1|^\nu} = \frac{1}{|\alpha+1|(|\alpha+1|-1)},$$

[29] Mellin, H., Om gammafunktionen. Öfversigt Akad. Stockholm, 1883, no. 5, pp. 3–20.

whence we conclude that for $|s+1|, |s+2|, \cdots$ all greater than unity,

$$|\omega^*(s)| < \sum_1^\infty \frac{1}{|s+\nu|(|s+\nu|-1)},$$

a series which we have already studied in § 5. When $|s|$ increases in the manner indicated there, $\omega^*(s) \to 0$, or

$$\lim_{|s|\to\infty} [\psi(s) - \log s] = 0, \tag{29}$$

which extends the domain of validity of (21). For $|s+1|, |s+2|, \cdots$ all greater than unity we find

$$\omega^*(s) = \sum_{\nu=2}^\infty \left(1 - \frac{1}{\nu}\right) \sum_{\mu=1}^\infty \frac{1}{(s+\mu)^\nu}. \tag{30}$$

On the other hand, the expansion

$$\frac{1}{\alpha} - \log\left(1 + \frac{1}{\alpha}\right) = \sum_2^\infty \frac{(-1)^\nu}{\nu} \frac{1}{\alpha^\nu}, \qquad (|\alpha| > 1),$$

gives

$$\omega^*(s) = \sum_{\nu=2}^\infty \frac{(-1)^\nu}{\nu} \sum_{\mu=0}^\infty \frac{1}{(s+\mu)^\nu}, \tag{31}$$

which is valid for $|s|, |s+1|, \cdots$ all greater than unity. Expansions more convenient than (30) and (31) for computing $\omega^*(s)$ numerically, will be derived in § 14.

Equation (29) may be used to determine approximately the roots of the transcendental equation $\psi(s) = 0$. According to (19), the imaginary part of $\psi(x + yi)$ equals

$$yi \cdot \sum_0^\infty \frac{1}{(x+\nu)^2 + y^2},$$

which cannot vanish unless $y = 0$, so that all the roots must be real. Furthermore it follows from (19''') that

$$\psi(x) - \psi(x') = (x - x') \sum_0^\infty \frac{1}{(x+\nu)(x'+\nu)},$$

and every term in the series is positive when x and x' are both in one of the intervals $(0 \cdots +\infty)$, $(-1 \cdots 0)$, $(-2 \cdots -1)$, $\cdots$; since $\psi(+\infty) = +\infty$, $\psi(-n \pm 0) = \mp\infty$, it is seen that $\psi(x)$ increases steadily from $-\infty$ to $+\infty$ in any one of the intervals in question, so that each interval contains one root only. The positive root lies between 1 and 2, since $\psi(1) = -C < 0$, $\psi(2) = 1 - C > 0$, and its value $1.46163\cdots$ was computed by Legendre. Denoting the negative root in the interval

$(-n \cdots -n+1)$ by $-n+x_n$, $(0 < x_n < 1)$, it follows from (22) that

$$\pi \cot \pi x_n = \psi(n+1-x_n) = \log n + \delta_n,$$

where, according to (29), $\delta_n \to 0$ for $n \to \infty$. Therefore

$$x_n = \frac{1}{\pi} \arctan \frac{\pi}{\log n + \delta_n},$$

whence we conclude that $x_n \log n \to 1$. For large values of n, the root in the interval $(-n \cdots -n+1)$ is thus approximately equal to

$$-n + \frac{1}{\log n},$$

a result due to Hermite.[30] Remembering that $\Gamma(x)$ never equals zero, we see at once that $\Gamma'(x) = \Gamma(x)\psi(x) = 0$ must have the same roots as $\psi(x) = 0$, which is of importance when inquiring how $\Gamma(x)$ varies for real values of x.

9. Evaluation of $\psi(x)$ for rational values of x. On account of (20), it is sufficient to evaluate

$$\psi\left(\frac{p}{q}\right) = -C + \sum_{\nu=0}^{\infty}\left(\frac{1}{\nu+1} - \frac{q}{p+\nu q}\right),$$

when p/q is a *proper* fraction (p and q integers, $q > p > 0$). Instead of this expansion we shall consider the power series

$$S(t) = \sum_{\nu=0}^{\infty}\left(\frac{1}{\nu+1} - \frac{q}{p+\nu q}\right) t^{p+\nu q},$$

which is convergent for $t = 1$. By Abel's theorem on power series,[31]

[30] Hermite, Ch., Sur l'intégrale Eulérienne de seconde espèce. Journal für Math., vol. 90 (1881), pp. 332–338.

[31] ["When $\sum_0^\infty a_\nu$ is convergent, the power series $F(t) = \sum_0^\infty a_\nu t^\nu$ approaches the limit $\sum_0^\infty a_\nu$ when t increases through real positive values toward unity." The proof commonly given is this: $\sum_0^\infty a_\nu$ being convergent, the remainder after n terms $r_n = \sum_n^\infty a_\nu \to 0$ for $n \to \infty$, and hence there exists a positive quantity A such that $|r_n| < A$ for $n = 0, 1, 2, \cdots$ and to any $\epsilon > 0$ but as small as we please, there corresponds an integer N such that $|r_n| < \epsilon$ for $n > N$. Obviously

$$a_\nu = r_\nu - r_{\nu+1},$$

and consequently for $0 < t < 1$

$$F(t) - \sum_0^\infty a_\nu = \sum_1^\infty a_\nu(t^\nu - 1) = \sum_1^\infty (r^\nu - r_{\nu+1})(t^\nu - 1)$$

$$= \lim_{n\to\infty} [(r_1 - r_2)(t-1) + (r_2 - r_3)(t^2 - 1) + \cdots + (r_n - r_{n+1})(t^n - 1)]$$

$$= \lim_{n\to\infty} [-r_1(1-t) - r_2(t-t^2) - \cdots - r_n(t^{n-1} - t^n)] + \lim_{n\to\infty} r_{n+1}(1-t^n),$$

$$\psi\left(\frac{p}{q}\right) = -C + \lim_{t \to 1} S(t).$$

We have, however, for $|t| < 1$,

$$\sum_0^\infty \frac{t^{p+\nu q}}{\nu + 1} = -t^{p-q} \log(1 - t^q),$$

and the problem thus reduces to the determination, for $|t| < 1$, of $q \sum_0^\infty \frac{t^{p+\nu q}}{p + \nu q}$, which consists of q times every qth term in the series

$$\sum_1^\infty \frac{t^\nu}{\nu} = -\log(1 - t).$$

But given the power series $F(t) = \sum_0^\infty c_\nu t^\nu$, it is well known that, for $\theta = e^{2\pi i/q}$,

$$\sum_{\mu=0}^{q-1} \theta^{-\mu p} F(t\theta^\mu) = q(c_p t^p + c_{p+q} t^{p+q} + c_{p+2q} t^{p+2q} + \cdots),$$

since

$$\sum_{\mu=0}^{q-1} \theta^{\mu(\nu-p)} = \frac{\theta^{q(\nu-p)} - 1}{\theta^{\nu-p} - 1} = \frac{1 - 1}{\theta^{\nu-p} - 1} = 0$$

when $\nu - p$ is not divisible by q, but

$$= \sum_{\mu=0}^{q-1} 1 = q$$

when $\nu - p$ is divisible by q. Hence we find, for $|t| < 1$,

$$S(t) = -t^{p-q} \log(1 - t^q) + \sum_{\mu=0}^{q-1} \theta^{-\mu p} \log(1 - t\theta^\mu)$$

---

and since $r_{n+1} \to 0$, we have

$$F(t) - \sum_0^\infty a_\nu = -\sum_1^\infty r_\nu(t^{\nu-1} - t^\nu) = -(1 - t)\sum_1^N r_\nu t^{\nu-1} - \sum_{N+1}^\infty r_\nu(t^{\nu-1} - t^\nu).$$

For $0 < t < 1$, it is clear that $1 - t > 0$, $t^{\nu-1} - t^\nu > 0$, whence

$$\begin{aligned}\left|F(t) - \sum_0^\infty a_\nu\right| &\leqq (1 - t)\sum_1^N |r_\nu| t^{\nu-1} + \sum_{N+1}^\infty |r_\nu| (t^{\nu-1} - t^\nu) \\ &< (1 - t)\sum_1^N A \cdot 1 + \sum_{N+1}^\infty \epsilon(t^{\nu-1} - t^\nu) \\ &= (1 - t)NA + \epsilon t^N < NA(1 - t) + \epsilon \\ &< 2\epsilon \text{ for } 0 < 1 - t < \epsilon/NA,\end{aligned}$$

so that $\left|F(t) - \sum_0^\infty a_\nu\right|$ can be made as small as we please by taking t positive and less than, but sufficiently close to, unity.]

$$= -t^{p-q}\log\frac{1-t^q}{1-t}+\sum_{\mu=1}^{q-1}\theta^{-\mu p}\log(1-t\theta^\mu)-(t^{p-q}-1)\log(1-t),$$

and letting t increase to unity along the real axis

$$\psi\left(\frac{p}{q}\right)=-C-\log q+\sum_{\mu=1}^{q-1}\theta^{-\mu p}\log(1-\theta^\mu).$$

Thus our problem is solved, and it only remains to transform the expression obtained. Replacing p by $q-p$, a new equation results, and adding it to the above, we find

$$\psi\left(\frac{p}{q}\right)+\psi\left(\frac{q-p}{q}\right)=-2C-2\log q+2\sum_{\mu=1}^{q-1}\cos\frac{2\pi\mu p}{q}\log(1-\theta^\mu).$$

Since $\psi(p/q)$ is real, the imaginary part of the right member vanishes identically, so that

$$\psi\left(\frac{p}{q}\right)+\psi\left(\frac{q-p}{q}\right)=-2C-2\log q+\sum_{\mu=1}^{q-1}\cos\frac{2\pi\mu p}{q}\log\left(2-2\cos\frac{2\pi\mu}{q}\right),$$

which, together with

$$\psi\left(\frac{p}{q}\right)-\psi\left(\frac{q-p}{q}\right)=-\pi\cot\frac{\pi p}{q},$$

gives

$$\psi\left(\frac{p}{q}\right)=-C-\frac{\pi}{2}\cot\frac{\pi p}{q}-\log q+\frac{1}{2}\sum_{\mu=1}^{q-1}\cos\frac{2\pi\mu p}{q}\log\left(2-2\cos\frac{2\pi\mu}{q}\right).$$

Since

$$\cos\frac{2\pi(q-\mu)}{q}=\cos\frac{2\pi\mu}{q}\qquad\text{and}\qquad\cos\frac{2\pi(q-\mu)p}{q}=\cos\frac{2\pi\mu p}{q},$$

the summation on the right side may be extended from $\mu = 1$ to $\mu = \dfrac{q-1}{2}$ or $\dfrac{q}{2}$ according as q is odd or even. The final result thus becomes

$$\psi\left(\frac{p}{q}\right)=-C-\frac{\pi}{2}\cot\frac{\pi p}{q}-\log q+\sum_{\mu=1}^{q'}{}'\cos\frac{2\pi\mu p}{q}\log\left(2-2\cos\frac{2\pi\mu}{q}\right),\tag{32}$$

where $q' = q/2$ or $(q-1)/2$ according as q is even or odd, and the accent to the right of the summation sign indicates that the term corresponding to $\mu = q/2$ in the first case should be divided by 2. The proof given here is clearer and much simpler than that of Gauss,[12] to whom is due the remarkable theorem expressed by (32). As special cases, we find

$\frac{p}{q}$	$\psi\left(\frac{p}{q}\right) + C$
$\frac{1}{2}$	$-2\log 2$
$\frac{1}{3}$	$-\frac{\pi}{2\sqrt{3}} \quad -\frac{3}{2}\log 3$
$\frac{2}{3}$	$+\frac{\pi}{2\sqrt{3}} \quad -\frac{3}{2}\log 3$
$\frac{1}{4}$	$-\frac{\pi}{2} \quad -3\log 2$
$\frac{3}{4}$	$+\frac{\pi}{2} \quad -3\log 2$
$\frac{1}{5}$	$-\frac{\pi}{2}\sqrt{1+\frac{2}{\sqrt{5}}} - \frac{5}{4}\log 5 - \frac{\sqrt{5}}{4}\log\frac{3+\sqrt{5}}{2}$
$\frac{2}{5}$	$-\frac{\pi}{2}\sqrt{1-\frac{2}{\sqrt{5}}} - \frac{5}{4}\log 5 + \frac{\sqrt{5}}{4}\log\frac{3+\sqrt{5}}{2}$
$\frac{3}{5}$	$+\frac{\pi}{2}\sqrt{1-\frac{2}{\sqrt{5}}} - \frac{5}{4}\log 5 + \frac{\sqrt{5}}{4}\log\frac{3+2\sqrt{5}}{2}$
$\frac{4}{5}$	$+\frac{\pi}{2}\sqrt{1+\frac{2}{\sqrt{5}}} - \frac{5}{4}\log 5 - \frac{\sqrt{5}}{4}\log\frac{3+2\sqrt{5}}{2}$

10. The functions $P(s)$ and $Q(s)$. In § 2 we saw that $\Gamma(s)$ is a single-valued analytic function of s, having poles of the first order at the points $0, -1, -2, \cdots$, but holomorphic for all other finite values of s. In other words, $\Gamma(s+t)$ may be expanded in a power series in t for $|t|$ sufficiently small, unless s is zero or a negative integer. In the latter case, however, we have for sufficiently small values of $|t|$

$$\Gamma(-\nu+t) = \frac{(-1)^\nu}{\nu!}t^{-1} + a_0 + a_1 t + \cdots,$$

since

$$\lim_{t\to 0} t\Gamma(-\nu+t) = \lim_{t\to 0}\frac{\Gamma(1+t)}{(t-\nu)(t-\nu+1)\cdots(t-1)} = \frac{(-1)^\nu}{\nu!}.$$

We are thus led to form the infinite series

$$P(s) = \sum_0^\infty \frac{(-1)^\nu}{\nu!(\nu+s)}, \tag{33}$$

which is uniformly convergent for all finite s and becomes infinite at the same points and in the same manner as $\Gamma(s)$. For when N is chosen so large that $N > |s|$, we have, for all $\nu > N$,

$$\left|\frac{(-1)^\nu}{\nu!(\nu+s)}\right| < \frac{1}{\nu!(\nu-N)},$$

which is independent of s and forms the general term of a convergent series. Hence it follows that $P(s)$ is a single-valued analytic function of s in the entire plane, which can become infinite at the points $0, -1, -2, \cdots$ only, but is holomorphic for all other finite s. For sufficiently small values of $|t|$, we have

$$P(-\nu+t) = \frac{(-1)^\nu}{\nu!}t^{-1} + \alpha_0 + \alpha_1 t + \cdots.$$

It is evident, therefore, that $\Gamma(s+t) - P(s+t)$ may be expanded in a power series in t for any finite s and for $|t|$ sufficiently small, and by a well-known theorem, $\Gamma(s) - P(s)$ is consequently expansible in a power series in s which converges for all finite s and is denoted by $Q(s)$. The result of the above investigation which is due to Prym[32] may be expressed by writing

$$\Gamma(s) = P(s) + Q(s), \tag{34}$$

where $P(s)$ is defined by (33) while $Q(s)$ is an *entire transcendental*[33] *function* of s. We have

$$P(s+1) - sP(s) = \sum_0^\infty \frac{(-1)^\nu(\nu+1)}{(\nu+1)!(\nu+s+1)} - 1 - \sum_0^\infty \frac{(-1)^{\nu+1}s}{(\nu+1)!(\nu+s+1)}$$

$$= -1 + \sum_0^\infty \frac{(-1)^\nu}{(\nu+1)!} = -e^{-1},$$

and $P(s)$ consequently satisfies the functional equation

$$P(s+1) = sP(s) - e^{-1}, \tag{35}$$

while $Q(s)$, on account of (34) and (4a), will satisfy the analogous equation

$$Q(s+1) = sQ(s) + e^{-1}. \tag{36}$$

When the distance of s from the nearest point of the negative real axis increases indefinitely, $P(s)$ approaches zero since, denoting this distance by ρ, we have

$$|P(s)| < \sum_0^\infty \frac{1}{\nu!} \cdot \frac{1}{\rho} = \frac{e}{\rho}.$$

Hence, a fortiori,

$$\lim_{n\to\infty} \frac{P(n+s)}{(n-1)!n^s} = 0, \tag{37}$$

[32] Prym, F., Zur Theorie der Gammafunktion. Journal für Math., vol. 82 (1876), pp. 165–172.

[33] [That $Q(s)$ is transcendental follows most easily from equation (36) below, for supposing $Q(s)$ to be a polynomial of degree m, the polynomial $sQ(s)$, of degree $m+1$, would equal the polynomial $Q(s+1) - e^{-1}$, of degree m, which is clearly impossible.]

and by (34) and (4b)

$$\lim_{n\to\infty} \frac{Q(n+s)}{(n-1)!n^s} = 1. \tag{38}$$

That equations (35) and (37) are sufficient to define $P(s)$ (Prym[32]), and consequently (36) and (38) sufficient to define $Q(s)$, is seen in the following way, which will also furnish a new expression for $P(s)$. From (35) we find upon division by $\Gamma(s+1)$

$$\frac{P(s)}{\Gamma(s)} - \frac{P(s+1)}{\Gamma(s+1)} = \frac{e^{-1}}{\Gamma(s+1)},$$

whence, replacing s by $s+1, s+2, \cdots, s+n-1$ and adding the resulting equations,

$$\frac{P(s)}{\Gamma(s)} - \frac{P(s+n)}{\Gamma(s+n)} = e^{-1}\sum_1^n \frac{1}{\Gamma(s+\nu)} \tag{39}$$

or

$$\begin{aligned} P(s+n) &= s(s+1)\cdots(s+n-1) \\ &\times\left[P(s) - e^{-1}\sum_1^n \frac{1}{s(s+1)\cdots(s+\nu-1)}\right], \end{aligned} \tag{39'}$$

and since

$$\lim_{n\to\infty}\frac{P(n+s)}{\Gamma(n+s)} = \lim_{n\to\infty}\frac{P(n+s)}{(n-1)!n^s} \times \lim_{n\to\infty}\frac{(n-1)!n^s}{\Gamma(n+s)} = 0\times 1 = 0,$$

(39) gives for $n\to\infty$

$$\frac{P(s)}{\Gamma(s)} = e^{-1}\sum_1^\infty \frac{1}{\Gamma(s+\nu)} \tag{40}$$

or

$$P(s) = e^{-1}\sum_1^\infty \frac{1}{s(s+1)\cdots(s+\nu-1)} \tag{40'}$$

which is due to [Legendre and was rediscovered by] Bourguet.[34] Making $P(s) = \Gamma(s) - Q(s)$ in (39), we find

$$\frac{Q(s+n)}{\Gamma(s+n)} - \frac{Q(s)}{\Gamma(s)} = e^{-1}\sum_1^n \frac{1}{\Gamma(s+\nu)} \tag{41}$$

or

$$\begin{aligned} Q(s+n) &= s(s+1)\cdots(s+n-1) \\ &\times\left[Q(s) + e^{-1}\sum_1^n \frac{1}{s(s+1)\cdots(s+\nu-1)}\right]. \end{aligned} \tag{41'}$$

---

[34] [Legendre, A. M., Exercices de calcul intégral, vol. 1 (Paris, 1811), pp. 339–343.] Bourguet, L., Développement en séries des intégrales Eulériennes. Annales de l'École Normale, ser. 2, vol. 10 (1881), pp. 175–233.

For $s \to 0$ it follows from (35) and (36) that $P(1) = 1 - e^{-1}$, $Q(1) = e^{-1}$, and for $s = 1$, $P(2) = P(1) - e^{-1} = 1 - 2e^{-1}$, $Q(2) = 2e^{-1}$, etc. It is easy to obtain general expressions for $P(n)$ and $Q(n)$ from (39′) and (41′); we find for $s \to 0$

$$\begin{aligned} P(n) &= (n-1)!\left[1 - e^{-1}\sum_0^{n-1}\frac{1}{\nu!}\right], \\ Q(n) &= (n-1)!e^{-1}\sum_0^{n-1}\frac{1}{\nu!}. \end{aligned} \tag{42}$$

From (36) it is seen that

$$Q(0) = \lim_{s\to 0}\frac{Q(1+s) - e^{-1}}{s} = \lim_{s\to 0}\frac{Q(1+s) - Q(1)}{s} = Q'(1),$$

which we shall learn to calculate in § 13. Making $s = -n$ in (41′) it is further seen that

$$\begin{aligned} Q(-n) = {} & \frac{(-1)^n}{n!}Q(0) + e^{-1} \\ & \times\left[\frac{1}{n} - \frac{1}{n(n-1)} + \frac{1}{n(n-1)(n-2)} - \cdots + \frac{(-1)^{n-1}}{n!}\right]. \end{aligned} \tag{43}$$

11. Solution of a functional equation. The form of equations (35) and (36) furnishes an occasion to consider the more general functional equation

$$f(s+1) = sf(s) - R(s), \tag{44}$$

where $R(s)$ is a polynomial in s. Let $S(s)$ be some definite solution, then $S(s+1) = sS(s) - R(s)$, and consequently

$$f(s+1) - S(s+1) = s[f(s) - S(s)],$$

or according to § 1, $f(s) - S(s) = p(s)\Gamma(s)$, where $p(s)$ is an arbitrary periodic function with the additive period unity. The *most general function* satisfying equation (44) is therefore

$$f(s) = S(s) + p(s)\Gamma(s).$$

To find a particular solution $S(s)$ we divide (44) by $\Gamma(s+1)$, whence

$$\frac{f(s)}{\Gamma(s)} - \frac{f(s+1)}{\Gamma(s+1)} = \frac{R(s)}{\Gamma(s+1)},$$

and consequently

$$\frac{f(s)}{\Gamma(s)} - \frac{f(s+n)}{\Gamma(s+n)} = \sum_1^n \frac{R(s+\nu-1)}{\Gamma(s+\nu)}.$$

In this equation, we now let n increase indefinitely, the resulting infinite

series converging rapidly as shown by (4b), and we obtain

$$\frac{f(s)}{\Gamma(s)} - \lim_{n\to\infty} \frac{f(s+n)}{\Gamma(s+n)} = \sum_1^\infty \frac{R(s+\nu-1)}{\Gamma(s+\nu)},$$

where the limit on the left must exist and be finite. For the particular solution $f(s) = S(s)$ we assume this limit to be zero,[35] or

$$\lim_{n\to\infty} \frac{S(s+n)}{\Gamma(s+n)} = \lim_{n\to\infty} \frac{S(s+n)}{(n-1)!n^s} = 0. \tag{45}$$

Hence we find

$$\frac{S(s)}{\Gamma(s)} = \sum_1^\infty \frac{R(s+\nu-1)}{\Gamma(s+\nu)}. \tag{46}$$

Any polynomial of degree m may be written in the form

$$R(s) = a_0 + a_1 s + a_2 s(s-1) + \cdots + a_m s(s-1)\cdots(s-m+1),$$

as is readily seen by the method of undetermined coefficients, and since we have

$$\frac{R(s)}{\Gamma(s+1)} = \frac{a_0}{\Gamma(s+1)} + \frac{a_1}{\Gamma(s)} + \frac{a_2}{\Gamma(s-1)} + \cdots + \frac{a_m}{\Gamma(s-m+1)},$$

it follows from (40) that

$$\sum_{\nu=1}^\infty \frac{R(s+\nu-1)}{\Gamma(s+\nu)} = e\sum_{\mu=0}^m a_\mu \frac{P(s-\mu)}{\Gamma(s-\mu)}.$$

But in consequence of (39), $P(s-\mu)/\Gamma(s-\mu)$ may be written

$$\frac{P(s-\mu)}{\Gamma(s-\mu)} = \frac{P(s)}{\Gamma(s)} + e^{-1}\sum_{\nu=1}^{\mu} \frac{1}{\Gamma(s-\mu+\nu)},$$

and it is therefore seen at once that

$$\sum_{\nu=1}^\infty \frac{R(s+\nu-1)}{\Gamma(s+\nu)} = \frac{eP(s)}{\Gamma(s)}\sum_{\mu=0}^m a_\mu + \sum_{\mu=1}^m a_\mu \sum_{\nu=1}^{\mu} \frac{1}{\Gamma(s-\mu+\nu)},$$

or by (46), upon multiplication by $\Gamma(s)$,

$$S(s) = eP(s)\sum_{\mu=0}^m a_\mu + \sum_{\mu=1}^m a_\mu + \sum_{\mu=2}^m a_\mu \sum_{\nu=1}^{\mu-1} (s-1)(s-2)\cdots(s-\mu+\nu), \tag{47}$$

whereby the proposed functional equation is completely solved. Lindhagen[36] has derived this solution in a somewhat more complicated manner and without giving it in explicit form.

---

[35] [Carrying out a suggestion by the author, the translation differs here slightly from the original, in which the limit (45) is assumed to be a constant K, not necessarily zero.]

[36] Lindhagen, A., Studier öfver Gammafunktionen. Thesis, Upsala, 1887.

It would lead us too far here to discuss the solution of the more general functional equation

$$f(s+1) = r(s)f(s) + R(s),$$

where $r(s)$ and $R(s)$ are rational functions, and regarding which the researches of Mellin[37] may be consulted.

12. Investigation of $P(s)$ for real values of s. It appears immediately on inspection of the expansion (40′) that $P(x)$ is *positive and decreasing for real, positive and increasing* values of x. When x falls within the interval $(0 \cdots 1)$, we therefore have $P(x+1) > P(2) = 1 - 2e^{-1}$, and by (35)

$$eP(x) = \frac{eP(x+1)+1}{x} > \frac{e-1}{x}, \qquad (0 < x < 1),$$

an inequality which will be used below. In general it follows from (39′) that

$$(\alpha) \qquad eP(x-1) = \frac{eP(x)+1}{x-1},$$

$$(\beta) \qquad eP(x-2) = \frac{eP(x)+x}{(x-1)(x-2)},$$

$$(\gamma) \qquad eP(x-4) = \frac{eP(x)+x+(x-1)(x-2)^2}{(x-1)(x-2)(x-3)(x-4)}.$$

For an x in the interval $(0 \cdots 1)$, the numerator on the right in (α) is positive, the denominator negative, and hence $P(x)$ is *negative in the interval* $(-1 \cdots 0)$; on the other hand, both numerator and denominator are positive in (β), and consequently $P(x)$ is *positive in the interval* $(-2 \cdots -1)$. To show that the same is true for (γ), we use the inequality given above, and obtain, the denominator being positive,

$$eP(x-4) > \frac{e-1+x^2+x(x-1)(x-2)^2}{x(x-1)(x-2)(x-3)(x-4)}.$$

Making $x = 1 - y$, $0 < y < 1$, the numerator of the last fraction equals

$$\begin{aligned} e - 1 + (1-y)^2 - y(1-y)(1+y)^2 &\geqq e - 1 + (1-y)^2 - \tfrac{1}{4}(1+y)^2 \\ &= e - \tfrac{7}{4}y - \tfrac{3}{4}y(1-y) \\ &\geqq e - \tfrac{7}{4} - \tfrac{3}{16} = e - \tfrac{31}{16} > 0, \end{aligned}$$

---

[37] Mellin, H., Zur Theorie der Gammafunktion. Acta Math., vol. 8 (1886), pp. 37–80. Über einen Zusammenhang zwischen gewissen linearen Differential- und Differenzengleichungen, ibid., vol. 9 (1887), pp. 137–166. [Zur Theorie der linearen Differenzengleichungen erster Ordnung, ibid., vol. 15 (1891), pp. 317–384. Über den Zusammenhang zwischen den linearen Differential- und Differenzengleichungen, ibid., vol. 25 (1902), pp. 139–164.]

since $y(1-y) = \frac{1}{4} - (\frac{1}{2} - y)^2 \leqq \frac{1}{4}$. Hence $P(x)$ is likewise positive in the interval $(-4 \cdots -3)$. From (β) it is seen immediately that when $P(x)$ is *negative for a negative value of* x, *so is also* $P(x-2)$, and consequently $P(x)$ is *negative in the intervals* $(-3 \cdots -2)$, $(-5 \cdots -4)$, $\cdots(-2n-1 \cdots -2n) \cdots$.

It only remains to investigate the intervals $(-6 \cdots -5)$, $(-8 \cdots -7)$, $\cdots$. When $\epsilon > 0$ but as small as we please, (33) gives

$$P(-2n+\epsilon) = \frac{1}{(2n)!\,\epsilon} + \cdots; \qquad P(-2n+1-\epsilon) = \frac{1}{(2n-1)!\,\epsilon} + \cdots.$$

$P(x)$ must therefore be *positive in the latter intervals, provided* x *is sufficiently near either end of the interval.* But we may show that $P(x)$ will also *assume negative values in these intervals.* We have by (γ) and (β)

$$eP(-\tfrac{11}{2}) = \frac{eP(-\frac{3}{2}) - \frac{257}{8}}{\frac{5}{2}\cdot\frac{7}{2}\cdot\frac{9}{2}\cdot\frac{11}{2}} = \frac{eP(\frac{1}{2}) - \frac{755}{32}}{\frac{1}{2}\cdot\frac{3}{2}\cdot\frac{5}{2}\cdot\frac{7}{2}\cdot\frac{9}{2}\cdot\frac{11}{2}},$$

and since, by (40′),

$$eP(\tfrac{1}{2}) = 2\left(1 + \frac{1}{\frac{3}{2}} + \frac{1}{\frac{3}{2}\cdot\frac{5}{2}} + \cdots\right) < 2e,$$

$P(-\frac{11}{2})$ is visibly negative, and hence also $P(-\frac{15}{2})$, $\cdots P(-2n+\frac{1}{2})$, $\cdots$. That the negative values which $P(x)$ thus assumes in the intervals $(-6 \cdots -5)$, $(-8 \cdots -7)$, $\cdots$, cannot be large, appears from the equation

$$P(x) - \frac{1}{ex} = \frac{P(x+1)}{x},$$

which shows that when $P(x)$, $P(x+1)$ and x are negative, we must have $P(x) - 1/ex > 0$ or $|P(x)| < 1/e\,|x|$.

Since $P(x)$ is a continuous function of x in all the intervals in question, it is clear that the equation $P(x) = 0$ must have at least one root in each of the intervals $(-\frac{11}{2} \cdots -5)$, $(-6 \cdots -\frac{11}{2})$, $\cdots$ $(-2n+\frac{1}{2} \cdots -2n+1)$, $(-2n \cdots -2n+\frac{1}{2})$, $\cdots$ but none in the remaining intervals (Bourguet[38]).

Note. It is a defect of the preceding investigation that it gives no information regarding the possibility of more than one root in the interval considered.[39] Bourguet[40] has attempted to show that the equation

[38] Bourguet, L., Sur la fonction Eulérienne, Comptes rendus, vol. 96 (1883), pp. 1307–1310, and Acta Math., vol. 2 (1883), pp. 296–298.

[39] [That each of the intervals in question contains exactly one root has been shown recently by C. N. Haskins, On the zeros of the function $P(x)$, complementary to the incomplete Gamma function, Trans. Am. Math. Soc., vol. 16 (1915), pp. 405–412. For a much simpler proof, which also shows that there are exactly four imaginary roots, see T. H. Gronwall, Sur les zéros des fonctions $P(z)$ et $Q(z)$ associées à la fonction gamma, Ann. de l'Ecole Normale, 1916.]

[40] Bourguet, L., Sur la théorie des intégrales Eulériennes, Comptes rendus, vol. 96 (1883), pp. 1487–1490.

$P(x) = 0$ cannot have more than 4 imaginary roots, but his proof is not correct. We may note in this connection that it is still undecided whether the equation $Q(s) = 0$ has any roots or not.[41, 42] That there are none when the real part of $s - 1$ is negative, was shown by Lindhagen.[36] We cannot enter further upon these questions which require more powerful methods than those at our disposal here, but on another occasion, the author proposes to examine the equation $P(s) = k$ for k real and show that for $k = 0$, the equation has exactly 4 imaginary roots.

13. Calculation of the coefficients in some power series. By (19) we have

$$\psi(1 + s) = - C + \sum_0^\infty \left(\frac{1}{\nu + 1} - \frac{1}{\nu + 1 + s}\right),$$

whence, expanding each term in powers of s for $|s| < 1$,

$$\psi(1 + s) = - S_1 + S_2 s - S_3 s^2 + S_4 s^3 - \cdots, \tag{48}$$

$$S_1 = C, \qquad S_n = 1 + \frac{1}{2^n} + \frac{1}{3^n} + \cdots, \qquad (n > 1). \tag{48'}$$

The coefficients S_2 to S_{35} were computed by Legendre[43] to 16 decimals (15 correct) and S_2 to S_{70} by Stieltjes[44] to 32 decimals (30 correct). To determine the coefficients in the expansion

$$\Gamma(1 + s) = 1 + a_1 s + a_2 s^2 + \cdots, \tag{49}$$

which will be convergent for $|s| < 1$, we recollect that

$$\Gamma'(1 + s) = \psi(1 + s)\Gamma(1 + s),$$

or

$$a_1 + 2a_2 s + 3a_3 s^2 + \cdots = (- S_1 + S_2 s - S_3 s^2 + \cdots)(1 + a_1 s + a_2 s^2 + \cdots)$$

whence, comparing the coefficients of s^n on both sides,

$$(n + 1)a_{n+1} = - S_1 a_n + S_2 a_{n-1} - \cdots + (- 1)^{n+1} S_{n+1}, \tag{49'}$$

[41] Later note. In a paper read before the Mathematical Society of Copenhagen in 1893, the author has given approximate expressions for the roots of $Q(s) = 0$. This investigation has not yet been published [but will appear in an early number of these *Annals*].

[42] [The existence of an infinity of zeros of $Q(s)$ may be proved in the following manner, which is not elementary, but has the advantage of brevity. Assuming $Q(s) = 0$ to have only a finite number of roots, $s_1, s_2, \cdots, s_n$ (or none), the only roots of $Q(s) = e^{-1}$ will be $1, s_1 + 1, s_2 + 1, \cdots, s_n + 1$ since, by (36′), $Q(s) = e^{-1}$ is equivalent to $(s - 1)Q(s - 1) = 0$. But by a theorem due to Picard, an entire function which takes each of *two* different values (here 0 and e^{-1}) only a *finite* number of times, is a polynomial. Since $Q(s)$ cannot be a polynomial,[33] it follows that the equation $Q(s) = 0$ has an infinity of roots.]

[43] Legendre, A. M., Traité des fonctions elliptiques et des intégrales Eulériennes, vol. 2 (Paris, 1826), p. 432.

[44] Stieltjes, T. J., Table des valeurs des sommes S_k, Acta Math., vol. 10 (1887), pp. 299–302.

which may be used to compute the coefficients a_n numerically. For $n = 1, 2, \cdots$ we find successively

$$a_1 = -S_1, \qquad a_2 = \tfrac{1}{2}(S_1^2 + S_2), \qquad a_3 = -\tfrac{1}{6}(S_1^3 + 3S_1S_2 + 2S_3), \quad \cdots,$$

which however are less suited to numerical computation than (49′). Writing $(-1)^n a_n'$ instead of a_n in (49′), the alternating signs disappear, and it is seen that if $a_1', a_2', \cdots, a_n'$ are positive, this must also be true of a'_{n+1}. But $a_1' = S_1 > 0$, hence all a_n' are positive, and a_n *has the sign of* $(-1)^n$. It is also easy to calculate the coefficients of

$$(50) \qquad P(1 + s) = b_0 + b_1 s + b_2 s^2 + \cdots, \qquad (|s| < 1),$$

by expanding each term of

$$P(1 + s) = \sum_0^\infty \frac{(-1)^\nu}{\nu!(\nu + 1 + s)}.$$

We obtain

$$(50') \qquad b_0 = 1 - e^{-1}, \qquad b_n = \sum_{\nu=0}^\infty \frac{(-1)^{\nu+n}}{\nu!(\nu + 1)^{n+1}},$$

and these expressions are quite convenient for numerical computation. The infinite series for b_n consists of numerically decreasing terms with alternating signs, and consequently b_n has the same sign as the first term of the series, or the sign of $(-1)^n$.

Since $Q(1 + s) = \Gamma(1 + s) - P(1 + s)$, the coefficients in the power series

$$(51) \qquad Q(1 + s) = c_0 + c_1 s + c_2 s^2 + \cdots,$$

which converges for all finite s, are given by

$$(51') \qquad c_0 = e^{-1}, \qquad c_n = a_n - b_n.$$

Since $c_1 = Q'(1)$, this also determines $Q(0)$, as we saw at the end of § 10. The calculation of the coefficients in the expansion

$$(52) \qquad \frac{1}{\Gamma(1 + s)} = 1 + d_1 s + d_2 s^2 + \cdots,$$

which will converge for all finite s, may be effected by comparison of the coefficients of s^n in

$$(1 + a_1 s + a_2 s^2 + \cdots)(1 + d_1 s + d_2 s^2 + \cdots) = 1,$$

whence

$$(52') \qquad d_n + d_{n-1}a_1 + d_{n-2}a_2 + \cdots + a_n = 0,$$

which gives, using the previous expressions for $a_1, a_2, a_3, \cdots,$

$$d_1 = S_1, \qquad d_2 = \tfrac{1}{2}(S_1^2 - S_2), \qquad d_3 = \tfrac{1}{6}(S_1^3 - 3S_1S_2 + 2S_3), \quad \cdots.$$

Having once calculated the coefficients $a_1, a_2, a_3, \cdots$, (52′) serves to determine $d_1, d_2, d_3, \cdots$. When an independent determination is preferred, we may proceed as follows. By Taylor's theorem and (52) we find

$$\frac{1}{\Gamma(1+s+t)} = \frac{1}{\Gamma(1+s) + t\Gamma'(1+s) + \cdots}$$
$$= 1 + d_1(s+t) + d_2(s+t)^2 + \cdots$$

and comparing the coefficients of t on both sides of the last equality sign,

$$-\frac{1}{\Gamma(1+s)} \cdot \frac{\Gamma'(1+s)}{\Gamma(1+s)} = -\frac{1}{\Gamma(1+s)} \cdot \psi(1+s) = d_1 + 2d_2 s + 3d_3 s^2 + \cdots,$$

or

$$(1 + d_1 s + d_2 s^2 + \cdots)(S_1 - S_2 s + S_3 s^2 - \cdots) = d_1 + 2d_2 s + 3d_3 s^2 + \cdots,$$

whence finally

$$(52'') \qquad (n+1)d_{n+1} = S_1 d_n - S_2 d_{n-1} + \cdots + (-1)^n S_{n+1},$$

which is of the same form as (49′) except that the signs of $S_1, S_2, \cdots$ are changed. The coefficients of the power series for $\Gamma(1+s)$ were first computed by Jeffery.[45] The coefficients in the expansions of $1/\Gamma(2+s) = 1/(1+s)\Gamma(1+s)$, $1/\Gamma(s)$, $e^{Cs}/\Gamma(2+s)$, $\Gamma(2+s) = (1+s)\Gamma(1+s)$, $Q(s)$ and $eP(s)/\Gamma(s)$ were computed by Bourguet[34, 46] to 16 decimals, of which the last is doubtful.

In the papers quoted, the general recurrent formulas are derived in a somewhat more complicated manner than above.

14. Expansions in series of factorials. In the following, we shall expand various of the functions introduced previously in series of the form

$$\Sigma \frac{c_\nu}{s(s+1)\cdots(s+\nu-1)}.$$

To simplify, we introduce the notation

$$(s)_n = s(s+1)\cdots(s+n-1)$$

when n is a positive integer, while for $n = 0$ we write $(s)_0 = 1$. Furthermore we agree not to let s equal a negative integer or zero. It is then seen that

$$\frac{(s')_\nu}{(s)_\nu} - \frac{(s')_{\nu+1}}{(s)_{\nu+1}} = \frac{(s')_\nu}{(s)_\nu}\left(1 - \frac{s'+\nu}{s+\nu}\right) = \frac{(s')_\nu}{(s)_{\nu+1}}(s - s'),$$

[45] Jeffery, H. M., On the derivatives of the Gamma function, Quarterly Journal of Math., vol. 6 (1864), pp. 82–108.

[46] Bourguet, L., Sur les intégrales Eulériennes et quelques autres fonctions uniformes, Acta Math., vol. 2 (1883), pp. 261–295.

whence for $| s - s' | > 0$, making $\nu = 0, 1, \cdots, n - 1$ and adding,

$$(\alpha) \qquad \frac{1}{s - s'}\left[1 - \frac{(s')_n}{(s)_n}\right] = \sum_0^{n-1} \frac{(s')_\nu}{(s)_{\nu+1}}.$$

Since $(s)_n = \Gamma(n + s)/\Gamma(s)$, and consequently

$$\frac{(s')_n}{(s)_n} = \frac{\Gamma(s)}{\Gamma(s')} \cdot \frac{\Gamma(s' + n)}{(n - 1)!n^{s'}} \cdot \frac{(n - 1)!n^s}{\Gamma(s + n)} \cdot n^{s'-s},$$

it is seen at once that, denoting the real part of s by $\Re(s)$,

$$\lim_{n\to\infty} \frac{(s')_n}{(s)_n} = \begin{cases} 0 \\ \text{indeterminate, as } \Re(s - s') \\ \infty \end{cases} \begin{cases} > 0 \\ = 0. \\ < 0 \end{cases}$$

On account of (α) we therefore have the expansion, convergent for $\Re(s - s') > 0$,

$$(\beta) \qquad \frac{1}{s - s'} = \sum_0^\infty \frac{(s')_\nu}{(s)_{\nu+1}} = \frac{1}{s} + \frac{s'}{s(s + 1)} + \frac{s'(s' + 1)}{s(s + 1)(s + 2)} + \cdots$$

which is due to Stirling.[21] For $\Re(s - s') = 0$ the series will oscillate (between finite limits) and diverge for $\Re(s - s') < 0$. We have excluded above the case $s' = s$, where it is seen at once that the series reduces to $\Sigma\, 1/(s + \nu)$ and consequently diverges.

It does not appear from the preceding that the series is uniformly convergent, but this will be shown presently to be true for all values of s and s' subject to the conditions: s' finite, $\Re(s - s') \geqq \epsilon$, $| s |, | s + 1 |, | s + 2 |, \cdots \geqq \rho$, where ϵ and ρ are positive, but as small as we please.

We suppose these conditions fulfilled and write $s = x + yi$, $s' = x' + y'i$. The positive quantity A may then be chosen so large that $A > | s' |$, and consequently $x' > -A$, $| y' | < A$. Now taking an integer N satisfying the inequality

$$N > \frac{1}{\epsilon} A^2 + A,$$

we have for $\nu \geqq N$

$$\nu + x' \geqq N + x' > A + x' > 0$$

and

$$| \nu + s' | - (\nu + x') < \frac{\epsilon}{2},$$

since

$$(\nu + x')\left(\sqrt{1 + \left(\frac{y'}{\nu + x'}\right)^2} - 1\right) < \frac{y'^2}{2(\nu + x')} < \frac{A^2}{2\left[\frac{1}{\epsilon}A^2 + (A + x')\right]} = \frac{\epsilon}{2}.$$

Therefore

$$|\nu + s| - |\nu + s'| > \frac{\epsilon}{2},$$

since $|\nu + s| \geqq \nu + x \geqq \nu + x' + \epsilon$. We now write

$$u_n = \frac{(s')_n}{(s)_{n+1}}, \qquad \left|\frac{u_n}{u_{n+1}}\right| = \left|\frac{s+n+1}{s'+n}\right|,$$

and the preceding inequality becomes, making $\nu = n + 1$,

$$|s' + n| \cdot \left|\frac{u_n}{u_{n+1}}\right| - |s' + n + 1| > \frac{\epsilon}{2},$$

or

$$|u_{n+1}| < \frac{2}{\epsilon}(|s' + n| \cdot |u_n| - |s' + n + 1| \cdot |u_{n+1}|),$$

whence

$$|u_{n+1}| + |u_{n+2}| + \cdots < \frac{2}{\epsilon}|s' + n| \cdot |u_n| = \frac{2}{\epsilon}\prod_0^n \left|\frac{s' + \nu}{s + \nu}\right|.$$

But we have for all values of ν, and in particular for $\nu < N$,

$$\left|\frac{s' + \nu}{s + \nu}\right| < \frac{A + \nu}{\rho},$$

and for $\nu \geqq N$,

$$\left|\frac{s' + \nu}{s + \nu}\right| < \frac{|s' + \nu|}{|s' + \nu| + \frac{\epsilon}{2}} < \frac{\nu + A}{\nu + A + \frac{\epsilon}{2}}.$$

Consequently (if $n > N$),

$$\frac{2}{\epsilon}\prod_0^n \left|\frac{s' + \nu}{s + \nu}\right| < \frac{2}{\epsilon}\prod_0^{N-1}\left(\frac{A + \nu}{\rho}\right)\prod_N^n\left(\frac{\nu + A}{\nu + A + \frac{\epsilon}{2}}\right)$$

$$< \frac{2A(A+1)\cdots(A+N-1)}{\epsilon\rho^N \prod\limits_N^n \left(1 + \frac{\epsilon}{2(\nu + A)}\right)},$$

which is *independent* of s and s' and converges to zero for n increasing. It is thus shown that Σu_n or the right side of (β) is ABSOLUTELY *and* UNIFORMLY *convergent when s' is finite, $\Re(s - s') \geqq \epsilon > 0$ and* $|s|$, $|s + 1|$, $|s + 2|$, $\cdots \geqq \rho > 0$.

On the basis of this result it would be easy to show that the boundaries of the regions of conditional as well as absolute convergence of a series of the form $\Sigma\, c_\nu/(s)_\nu$ are straight lines perpendicular to the real axis, and it is also possible to state the necessary and sufficient conditions for the expansibility of a function in such a series. As we shall have no occasion

to use these results in the present paper, I shall not go beyond this indication.[47]

Assuming $|s| > |s'|$ and $\Re(s - s') > 0$, (β) may also be written

$$(\beta') \qquad \sum_0^\infty \frac{s'^\nu}{s^{\nu+1}} = \sum_0^\infty \frac{(s')_\nu}{(s)_{\nu+1}}.$$

The right hand series being uniformly convergent, it is permissible to expand in powers of s' and compare with the left side. Writing

$$(s')_n = C_0^n s'^n + C_1^n s'^{n-1} + \cdots + C_{n-1}^n s', \qquad C_0^n = 1,$$

where the first coefficients C_ν^n are readily calculated,[48] we find

$$(\gamma) \qquad \frac{1}{s^{\nu+1}} = \sum_{\mu=0}^\infty \frac{C_\mu^{\mu+\nu}}{(s)_{\mu+\nu+1}},$$

which will converge absolutely for $\Re(s) > 0$.*

[47] [An excellent account of this general theory is given by E. Landau, Über die Grundlagen der Theorie der Fakultätenreihen, Sitzungsber. Akad. München, vol. 36 (1906), pp. 151–218.]

[48] [Using, for instance, the identity $(s')_n = (s' + n - 1)(s')_{n-1}$, which gives the recurrent formula

$$C_\nu^n = C_\nu^{n-1} + (n-1)C_{\nu-1}^{n-1}.]$$

* {We will give the following direct proof that (γ) converges for $\Re(s) > 0$. The ratio of consecutive terms in the series is

$$\frac{C_\mu^{\nu+\mu}}{C_{\mu-1}^{\nu+\mu-1}(\mu + s + \nu)} = \frac{\mu + (\nu - 1) + C_\mu^{\nu+\mu-1}/C_{\mu-1}^{\nu+\mu-1}}{\mu + (s + \nu)}.$$

But, as we shall prove, $\lim_{\mu\to\infty} C_\mu^{\nu+\mu-1}/C_{\mu-1}^{\nu+\mu-1} = 0$. Hence the series will converge (and that absolutely) if $\Re(s) + \nu - (\nu - 1) > 1$, that is, if $\Re(s) > 0$.

That $\lim_{\mu\to\infty} C_\mu^{\nu+\mu-1}/C_{\mu-1}^{\nu+\mu-1} = 0$ may be shown as follows: $C_\mu^{\nu+\mu-1}$ and $C_{\mu-1}^{\nu+\mu-1}$ are the sums of all the products of the numbers $1, 2, \cdots, \nu + \mu - 2$, taken μ at a time, and $\mu - 1$ at a time, respectively. Hence $C_\mu^{\nu+\mu-1}$ may be written

$$C_\mu^{\nu+\mu-1} = \Sigma n_1 n_2 \cdots n_\mu, \qquad 1 \leqq n_1 < n_2 < \cdots < n_\mu \leqq \nu + \mu - 2.$$

To each term $n_1 n_2 \cdots n_\mu$ in this sum there corresponds a set of μ products in $C_{\mu-1}^{\nu+\mu-1}$ whose sum is

$$n_1 n_2 \cdots n_\mu \left(\frac{1}{n_1} + \frac{1}{n_2} + \cdots + \frac{1}{n_\mu}\right).$$

But if we form this sum for each term in $\Sigma n_1 n_2 \cdots n^\mu$ in turn, each product in $C_{\mu-1}^{\nu+\mu-1}$ will be taken $\nu - 1$ times, since between the number of products in $C_\mu^{\nu+\mu-1}$ and that in $C_{\mu-1}^{\nu+\mu-1}$ the relation $\mu\binom{\nu+\mu-2}{\mu} = (\nu - 1)\binom{\nu+\mu-2}{\mu-1}$ exists, where $\binom{n}{r}$ denotes the number of combinations of n letters taken r at a time. We therefore have

$$\frac{C_\mu^{\nu+\mu-1}}{(\nu - 1)C_{\mu-1}^{\nu+\mu-1}} = \frac{\Sigma n_1 n_2 \cdots n_\mu}{\Sigma n_1 n_2 \cdots n_\mu \left(\frac{1}{n_1} + \frac{1}{n_2} + \cdots + \frac{1}{n_\mu}\right)} < \frac{1}{\frac{1}{\nu - 1} + \frac{1}{\nu} + \cdots + \frac{1}{\nu + \mu - 2}}.$$

But the limit, as μ approaches ∞, of the last member of this inequality is 0. The same must therefore be true of $C_\mu^{\nu+\mu-1}/C_{\mu-1}^{\nu+\mu-1}$.}

When we have a series

$$f(s) = \sum_0^\infty \frac{a_\nu}{s^{\nu+1}}$$

converging absolutely for $|s| > r$, this may be transformed in a series of factorials with the aid of (γ). We then obtain a double series, absolutely convergent for $\Re(s) > r$, which becomes, by combining the coefficients (Stirling[21]),

$$(\delta) \qquad f(s) = \frac{a_0}{(s)_1} + \frac{a_1}{(s)_2} + \frac{a_2 + a_1}{(s)_3} + \frac{a_3 + 3a_2 + 2a_1}{(s)_4} + \cdots,$$

where the general term is

$$(\epsilon) \qquad \frac{A_\nu}{(s)_{\nu+1}} = \frac{a_\nu + C_1^\nu a_{\nu-1} + C_2^\nu a_{\nu-2} + \cdots + C_{\nu-1}^\nu a_1}{(s)_{\nu+1}}$$

or symbolically

$$\frac{a(a+1)\cdots(a+\nu-1)}{(s)_{\nu+1}},$$

which could also be derived directly from (β') by symbolic calculation. A direct proof that (δ) is *absolutely and uniformly* convergent for

$$\Re(s) \geqq r + \epsilon,$$

where ϵ is positive but arbitrarily small, is reached by observing that, since $\Sigma\, a_\nu/s^{\nu+1}$ converges for $|s| = r + \epsilon/2$, there exists a finite and positive M such that $|a_\nu| < M(r + \epsilon/2)^\nu$ for $\nu = 0, 1, 2, \cdots$. Hence

$$\left|\frac{A_\nu}{(s)_{\nu+1}}\right| < M \cdot \frac{\left(r + \frac{\epsilon}{2}\right)_\nu}{|(s)_{\nu+1}|},$$

and the theorem follows by comparison with the series (β) in which we make $s' = r + \epsilon/2$.

The identity

$$\frac{1}{(s+\nu)_m} - \frac{1}{(s+\nu+1)_m} = \frac{m}{(s+\nu)_{m+1}}$$

gives

$$(\zeta) \qquad \sum_{\nu=0}^{n-1} \frac{1}{(s+\nu)_{m+1}} = \frac{1}{m}\left[\frac{1}{(s)_m} - \frac{1}{(s+n)_m}\right],$$

whence we find for $m \geqq 1$ and $n \to \infty$ (Stirling[21])

$$(\eta) \qquad \sum_{\nu=0}^{\infty} \frac{1}{(s+\nu)_{m+1}} = \frac{1}{m}\frac{1}{(s)_m}.$$

It is readily seen that this series is absolutely and uniformly convergent for $\Re(s) \geqq -A$ and $|s|, |s+1|, |s+2|, \cdots \geqq \rho$, where A is finite and $\rho > 0$ but as small as we please.

Applying the transformation (ζ) to the series (δ), it is seen that, making $a_0 = 0$;

$$f(s) + f(s+1) + \cdots + f(s+n-1) = \sum_{\nu=1}^{\infty} \frac{A_\nu}{\nu}\left[\frac{1}{(s)_\nu} - \frac{1}{(s+n)_\nu}\right].$$

We have, however,

$$\frac{A_\nu}{\nu(s+n)_\nu} = \left(1 + \frac{s}{\nu}\right)\frac{s}{s+n}\,\frac{(s+1)_{\nu-1}}{(s+n+1)_{\nu-1}}\cdot\frac{A_\nu}{(s)_{\nu+1}};$$

here $\Re(s) > r$, s is finite, $|s+\mu|/|s+n+\mu| < 1$ for $\mu = 1, 2, \cdots$ so that

$$\frac{(s+1)_{\nu-1}}{(s+n+1)_{\nu-1}} < 1,$$

and consequently

$$\left|\frac{A_\nu}{\nu(s+n)_\nu}\right| < \left|1 + \frac{s}{\nu}\right|\cdot\frac{|s|}{|s+n|}\cdot\left|\frac{A_\nu}{(s)_{\nu+1}}\right|.$$

Since $\Sigma\,|A_\nu/(s)_{\nu+1}|$ is convergent, $|1 + s/\nu|$ finite, and $|s|/|s+n| \to 0$ as $n \to \infty$, we therefore find

$$\lim_{n\to\infty} \sum_{\nu=1}^{\infty} \frac{A_\nu}{\nu(s+n)_\nu} = 0,$$

whence finally

$$(\vartheta) \qquad f(s) + f(s+1) + f(s+2) + \cdots = \sum_{\nu=1}^{\infty} \frac{A_\nu}{\nu(s)_\nu}$$

for $\Re(s) > r$, which is also due to Stirling.[21, 49]

These formulas may be applied immediately to the calculation of $\omega(s)$. It will be remembered from §§ 4 and 5 that

$$\omega(s) = \sum_0^{\infty}\left[(s+\nu+\tfrac{1}{2})\log\left(1 + \frac{1}{s+\nu}\right) - 1\right],$$

and that for $|s| > 1$

$$(s+\tfrac{1}{2})\log\left(1+\frac{1}{s}\right) - 1 = \sum_1^{\infty}\frac{(-1)_\nu^{\nu+1}}{2(\nu+1)(\nu+2)s^{\nu+1}};$$

consequently, by (δ) and (ϵ),

$$(s+\tfrac{1}{2})\log\left(1+\frac{1}{s}\right) - 1 = \sum_1^{\infty}\frac{k_\nu}{(s)_{\nu+1}},$$

$$k_\nu = \frac{(-1)^{\nu+1}\nu}{2(\nu+1)(\nu+2)} + \frac{(-1)^\nu(\nu-1)}{2\nu(\nu+1)}C_1^\nu + \cdots + \tfrac{1}{12}C_{\nu-1}^\nu,$$

[49] [The series to the right in (ϑ) is absolutely and uniformly convergent for $\Re(s) \geqq r + \epsilon$. For as shown above, $|A_\nu| < M(r+\epsilon/2)_\nu$, whence

$$\left|\frac{A_\nu}{\nu(s)_\nu}\right| < M\frac{r+\frac{\epsilon}{2}+\nu-1}{\nu}\,\frac{\left(r+\frac{\epsilon}{2}\right)_{\nu-1}}{(s)_\nu} < M\left(r+\frac{\epsilon}{2}+1\right)\frac{\left(r+\frac{\epsilon}{2}\right)_{\nu-1}}{(s)_\nu},$$

and our statement is proved by comparison with the series (β) in which we make $s' = r + \epsilon/2$.

and from (ϑ) we finally obtain

(53) $$\omega(s) = \sum_{1}^{\infty} \frac{k_\nu}{\nu(s)_\nu}$$

which converges absolutely for $\Re(s) > 1$[50] and is due to Binet.[24] It is evident that $| k_\nu | < \frac{1}{12}(1)_\nu = \nu!/12$.

[50] [The above analysis does not give the complete domain of convergence. Actually (53) converges absolutely and uniformly for $\Re(s) \geqq \epsilon > 0$ (and consequently equals the holomorphic function $\omega(s)$ not only for $\Re(s) > 1$ but for $\Re(s) > 0$), but is divergent when $\Re(s) < 0$. For observing that $\mu/2(\mu+1)(\mu+2) = 1/(\mu+2) - 1/2(\mu+1)$, we have

$$k_\nu = - \sum_{\mu=1}^{\nu} \left(\frac{(-1)^{\mu+2}}{\mu+2} + \frac{(-1)^{\mu+1}}{2(\mu+1)} \right) C^\nu_{\nu-\mu} = - \int_0^{-1} \sum_{\mu=1}^{\nu} (s^{\mu+1} + \tfrac{1}{2}s^\mu) C^\nu_{\nu-\mu}\, ds,$$

or by reference to the identity defining the coefficients $C^\nu_{\nu-\mu}$,

$$-k_\nu = \int_0^{-1} (s + \tfrac{1}{2})(s)_\nu ds.$$

Replacing s by $-s$ in the integral, we obtain, for $\nu > 2$,

$$-k_\nu = \int_0^1 (\tfrac{1}{2} - s)s(1-s)(2-s)\cdots(\nu-1-s)ds.$$

But for $0 < s < 1$, we have $| 1/2 - s | < 1, s < 1, 1 - s < 1, 2 - s < 2, \cdots, \nu - 1 - s < \nu - 1$, so that

$$| k_\nu | < \int_0^1 1 \cdot 2 \cdot \cdots (\nu - 1) ds = (\nu - 1)!,$$

and for $\Re(s) \geqq \epsilon$ and $\nu > 2$,

$$\left| \frac{k_\nu}{\nu(s)_\nu} \right| < \frac{(\nu-1)(1)_{\nu-2}}{\nu\,|s|\,|(s+1)_{\nu-1}|} < \frac{1}{\epsilon} \cdot \frac{(1)_{\nu-2}}{|(s+1)_{\nu-1}|};$$

in consequence of what was proved above for the series (β), it follows that (53) is absolutely and uniformly convergent for $\Re(s+1) \geqq 1 + \epsilon$ or $\Re(s) \geqq \epsilon$. Now transform the expression for $-k_\nu$ by decomposing the integral in two with the limits $0, \frac{1}{2}$ and $\frac{1}{2}, 1$ respectively, and replacing s by $1 - s$ in the latter. Thus

$$-k_\nu = \int_0^{1/2} (\tfrac{1}{2} - s)s(1-s)[(2-s)\cdots(\nu-1-s) - (1+s)\cdots(\nu-2+s)]ds,$$

where the integrand is evidently positive, and therefore

$$| k_\nu | > \int_0^\delta (\tfrac{1}{2} - s)s(1-s)[(2-s)\cdots(\nu-1-s) - (1+s)\cdots(\nu-2+s)]ds$$

for a positive δ less than $\frac{1}{2}$ and as small as we please. But for $0 < s < \delta < \frac{1}{3}$ we have $1 + s < \frac{4}{5}(2 - s)$, $2 + s < 3 - s$, $\cdots$, $\nu - 2 + s < \nu - 1 - s$, so that $(1+s)\cdots(\nu-2+s) < \frac{4}{5}(2-s)\cdots(\nu-1-s)$, and consequently

$$\left| \frac{k_\nu}{\nu} \right| > \frac{1}{5} \int_0^\delta (\tfrac{1}{2} - s)s(1-s)(2-s)\cdots(\nu-2-s)\frac{\nu-1-s}{\nu} ds$$

$$> \frac{1}{5} \int_0^\delta (\tfrac{1}{2} - \tfrac{1}{3})s(1-\delta)(2-\delta)\cdots(\nu-2-\delta) \cdot \tfrac{1}{2} ds$$

$$= \frac{1}{120} \delta^2(1-\delta)(2-\delta)\cdots(\nu-2-\delta) = \frac{\delta}{72} |(-\delta)_{\nu-1}|.$$

Hence $\left| \frac{k_\nu}{\nu(s)_\nu} \right| > \frac{\delta}{120} \left| \frac{(-\delta)_{\nu-1}}{(s)_\nu} \right| \to \infty$ as $\nu \to \infty$ when $\Re(s) < -\delta$, so that (53) diverges when $\Re(s) < -\delta$].

The first coefficients are $k_1 = \frac{1}{12}, k_2 = 0, k_3 = -\frac{1}{120}, k_4 = -\frac{1}{30}, \cdots$. When s is large, the computation of only a few terms in (53) gives a very close approximation.

In the same way we obtain from

$$(s+\tfrac{1}{2})\log\left(1+\frac{1}{s}\right) - 1 = \sum_1^\infty \frac{\nu}{2(\nu+1)(\nu+2)(s+1)^{\nu+1}},$$

which is valid for $|s+1| > 1$,

$$(s+\tfrac{1}{2})\log\left(1+\frac{1}{s}\right) - 1 = \sum_1^\infty \frac{k_\nu}{(s+1)_{\nu+1}},$$

$$k_\nu = \frac{\nu}{2(\nu+1)(\nu+2)} + \frac{\nu-1}{2\nu(\nu+1)}C_1^\nu + \cdots + \tfrac{1}{12}C_{\nu-1}^\nu,$$

and

(54)
$$\omega(s) = \sum_1^\infty \frac{k_\nu}{\nu(s+1)_\nu},$$

which is also due to Binet[24] and converges absolutely and uniformly when $\Re(s) \geqq \epsilon > 0$.[51] The first coefficients are $k_1 = \frac{1}{12}$, $k_2 = \frac{1}{6}$, $k_3 = \frac{59}{120}$, $k_4 = \frac{29}{15}$ $\cdots$. Making $s = n$, a positive integer, in (53) and (54), we obtain [see (14)] the convergent series

$$\log(n!) - (n+\tfrac{1}{2})\log n + n - \log\sqrt{2\pi}$$

$$= \frac{1}{12n} - \frac{1}{360n(n+1)(n+2)} - \frac{1}{120n(n+1)(n+2)(n+3)} - \cdots$$

$$= \frac{1}{12(n+1)} + \frac{1}{12(n+1)(n+2)} + \frac{59}{360(n+1)(n+2)(n+3)} + \frac{29}{60(n+1)(n+2)(n+3)(n+4)} + \cdots.$$

Remembering that (§ 8)

$$\omega^*(s) = \sum_0^\infty\left[\frac{1}{s+\nu} - \log\left(1+\frac{1}{s+\nu}\right)\right],$$

and that for $|s| > 1$, on account of (δ) and (ϵ),

$$\frac{1}{s} - \log\left(1+\frac{1}{s}\right) = \sum_1^\infty \frac{(-1)^{\nu+1}}{(\nu+1)s^{\nu+1}} = \sum_1^\infty \frac{k_\nu}{(s)_{\nu+1}},$$

where

$$k_\nu = \frac{(-1)^{\nu+1}}{\nu+1} + \frac{(-1)^\nu}{\nu}C_1^\nu + \cdots + \tfrac{1}{2}C_{\nu-1}^\nu,$$

we find by (ϑ)

(55)
$$\omega^*(s) = \sum_1^\infty \frac{k_\nu}{\nu(s)_\nu},$$

---

[51] [Using the method of [50], the series is readily seen to diverge for $\Re(s) < 0$.]

which will converge absolutely for $\Re(s) > 1$,[52] and in which the first coefficients are $k_1 = \frac{1}{2}$, $k_2 = \frac{1}{6}$, $k_3 = \frac{1}{4}$, $k_4 = \frac{19}{30}$, $\cdots$. Similarly, using the expansion

$$\frac{1}{s} - \log\left(1 + \frac{1}{s}\right) = \sum_1^\infty \frac{\nu}{(\nu+1)(s+1)^{\nu+1}} = \sum_1^\infty \frac{k_\nu}{(s+1)_{\nu+1}},$$

which converges for $\Re(s) > 0$, we find

$$k_\nu = \frac{\nu}{\nu+1} + \frac{\nu-1}{\nu} C_1^\nu + \cdots + \tfrac{1}{2} C_{\nu-1}^\nu$$

and

$$\omega^*(s) = \sum_1^\infty \frac{k_\nu}{\nu(s+1)_\nu}, \tag{56}$$

which is absolutely and uniformly convergent for $\Re(s) \geqq \epsilon > 0$.[53] The first coefficients are $k_1 = \frac{1}{2}$, $k_2 = \frac{7}{6}$, $k_3 = \frac{15}{4}$, $k_4 = \frac{469}{30}$, $\cdots$. As special cases of (55) and (56), we note the convergent series

$$\begin{aligned}
1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \frac{1}{n-1} - \log n - C \\
&= -\frac{1}{2n} - \frac{1}{12n(n+1)} - \frac{1}{12n(n+1)(n+2)} \\
&\qquad - \frac{19}{120n(n+1)(n+2)(n+3)} - \cdots \\
&= -\frac{1}{2(n+1)} - \frac{7}{12(n+1)(n+2)} - \frac{5}{4(n+1)(n+2)(n+3)} \\
&\qquad - \frac{469}{120(n+1)(n+2)(n+3)(n+4)} - \cdots.
\end{aligned}$$

We shall finally give an application to the function $\psi(s)$. By (19‴) we have

$$\psi(s) - \psi(s-s') = \sum_0^\infty \left(\frac{1}{s-s'+\nu} - \frac{1}{s+\nu}\right),$$

and since for $\Re(s-s') > 0$

$$\frac{1}{s-s'} - \frac{1}{s} = \sum_1^\infty \frac{(s')_\nu}{(s)_{\nu+1}},$$

---

[52] [Actually the series converges absolutely and uniformly for $\Re(s) \geqq \epsilon > 0$, but diverges for $\Re(s) < 0$. For we have, by the method of [50],

$$k_\nu = \int_0^1 s(1-s)(2-s)\cdots(\nu-1-s)ds,$$

whence $0 < k_\nu < (\nu-1)!$, and for $0 < \delta < 1$,

$$\frac{k_\nu}{\nu} > \int_0^\delta s(1-\delta)(2-\delta)\cdots(\nu-2-\delta)\frac{\nu-1-\delta}{\nu}ds > \frac{\delta}{4}|(-\delta)_{\nu-1}|.]$$

[53] [But divergent for $\Re(s) < 0$, since $k_\nu > \frac{1}{2} + \frac{1}{2}C_1^\nu + \cdots + \frac{1}{2}C_{\nu-1}^\nu = \frac{1}{2}(\nu-1)!$]

it follows from (ϑ) that

$$\psi(s) - \psi(s - s') = \sum_1^\infty \frac{(s')_\nu}{\nu(s)_\nu},$$

or replacing s' by $-s'$,

$$\psi(s + s') - \psi(s) = \frac{s'}{s} - \frac{s'(s' - 1)}{2s(s + 1)} + \frac{s'(s' - 1)(s' - 2)}{3s(s + 1)(s + 2)} - \cdots, \tag{57}$$

which converges absolutely and uniformly for $\Re(s + s') \geqq \epsilon > 0$.[54] Assuming s' to be a positive integer, we obtain the finite series

$$\frac{1}{s} + \frac{1}{s + 1} + \cdots + \frac{1}{s + n - 1} = \frac{n}{s} - \frac{n(n - 1)}{2s(s + 1)} + \cdots.$$

15. Numerical computation of the gamma function.[55] Let us now consider a practical method of computing $\Gamma(s)$ for real values of s. It is easily shown that $\Gamma(s)$ is known when its values in *any* interval of length $\frac{1}{2}$ are known.

From (5) we have

$$L(s) = \log \Gamma(s) = -Cs - \log s + \sum_{n=1}^\infty \left\{ \frac{s}{n} - \log\left(1 + \frac{s}{n}\right)\right\}. \tag{58}$$

It is seen that it is permissible to differentiate the series on the right termwise; so that

$$L'(s) = -C - \frac{1}{s} + \sum_1^\infty \left\{\frac{1}{n} - \frac{1}{s + n}\right\}.$$

This may be written

$$L'(s) = -C + \sum_1^\infty \left\{\frac{1}{n} - \frac{1}{s + n - 1}\right\}.$$

These relations hold for $s \neq 0, -1, -2, \cdots$. In general we have

$$L^{(r)}(s) = (-1)^r (r - 1)! \sum_1^\infty \frac{1}{(s + n - 1)^r}, \qquad r > 1. \tag{59}$$

In particular

$$L'(1) = -C, \tag{60}$$

$$L^{(r)}(1) = (-1)^r(r - 1)! \Sigma \frac{1}{n^r} = (-1)^r (r - 1)! S_r, \tag{61}$$

where S_r has the meaning of (48′).

Using Taylor's development about $s = 1$, we have

$$\log \Gamma(s) = L(s) = L(1) + \frac{s - 1}{1!} L'(1) + \frac{(s - 1)^2}{2!} L''(1) + \cdots;$$

[54] [But evidently diverges for $\Re(s + s') < 0$.]

[55] This section has been added in the translation.

or setting $s = 1 + x$ and taking account of (60) and (61)

$$\log \Gamma(1 + x) = - Cx + \sum_{n=2}^{\infty} \frac{(-1)^n}{n} S_n x^n. \tag{62}$$

We show now that this relation is valid for $-\frac{1}{2} \leq x \leq \frac{1}{2}$, by proving that

$$R_m = \frac{x^m}{m!} L^{(m)}(1 + \theta x), \qquad 0 < \theta < 1$$

converges to 0, as $m \to \infty$.

For, if $-\frac{1}{2} \leq x \leq \frac{1}{2}$, then

$$|R_m| \leq \left\{ \left| \frac{x}{1 + \theta x} \right|^m + \sum_2^{\infty} \frac{x^m}{|n + \theta x|^m} \right\} \cdot \frac{1}{m} < \left\{ 1 + \sum_2^{\infty} \frac{1}{2^m (n-1)^m} \right\} \frac{1}{m} \to 0.$$

As a matter of fact the region of convergence could be shown to be $-1 < x \leq 1$, but for our purposes it suffices to know that the series converges for $-\frac{1}{2} \leq x \leq \frac{1}{2}$. We have for any x in $(-\frac{1}{2}, \frac{1}{2})$

$$\log (1 + x) = x - \sum_2^{\infty} (-1)^n \frac{x^n}{n}.$$

This on adding and subtracting from (62) gives

$$\log \Gamma(1 + x) = - \log (1 + x) + (1 - C)x + \sum_2^{\infty} (-1)^n (S_n - 1) \frac{x^n}{n}.$$

Changing here x into $-x$ gives

$$\log \Gamma(1 - x) = - \log (1 - x) - (1 - C)x + \sum_2^{\infty} (S_n - 1) \frac{x^n}{n}.$$

Subtracting this from the foregoing gives

$$\log \Gamma(1 + x) - \log \Gamma(1 - x)$$
$$= - \log \frac{1 + x}{1 - x} + 2(1 - C)x - \sum_1^{\infty} \frac{x^{2m+1}}{2m + 1} (S_{2m+1} - 1).$$

From (6), substituting therein $\Gamma(1 + s)/s$ for $\Gamma(s)$ and setting $s = x$,

$$\log \Gamma(1 + x) + \log \Gamma(1 - x) = \log \frac{\pi x}{\sin \pi x}.$$

This with the preceding relation gives

$$\log \Gamma(1 + x)$$
$$= (1 - C)x + \tfrac{1}{2} \log \frac{\pi x}{\sin \pi x} - \tfrac{1}{2} \log \frac{1 + x}{1 - x} - \frac{1}{2} \sum_1^{\infty} (S_{2m+1} - 1) \frac{x^{2m+1}}{2m + 1}. \tag{63}$$

This series is due to Legendre. It converges rapidly for $0 \leq x \leq \frac{1}{2}$ and enables one to compute $\Gamma(s)$ in the interval $1 \leq s \leq \frac{3}{2}$. The value of $\Gamma(s)$ for other real values of s may then be obtained as already observed.

ON THE WRONSKIAN TEST FOR LINEAR DEPENDENCE.

By Maxime Bôcher.

It is well known that the identical vanishing of the Wronskian, while a necessary condition for the linear dependence of functions of one variable* for which the necessary derivatives exist, is not a sufficient condition. It is, however, true that:

I. If $u_1(x), \cdots u_n(x)$ are throughout (a, b) analytic functions whose Wronskian $W(u_1, \cdots u_n)$ vanishes identically, then $u_1, \cdots u_n$ are linearly dependent throughout (a, b).

II. If $u_1(x), \cdots u_n(x)$ satisfy a homogeneous linear differential equation

$$\frac{d^k u}{dx^k} + p_1 \frac{d^{k-1}u}{dx^{k-1}} + \cdots + p_k u = 0 \tag{1}$$

at every point of an interval (a, b) throughout which $p_1, \cdots p_k$ are continuous, and if $W(u_1, \cdots u_n) \equiv 0$ in (a, b), then $u_1, \cdots u_n$ are linearly dependent in (a, b).

These two well-known theorems would probably be very nearly (if not quite) sufficient for all applications that have so far been made.

There are, however, many other cases which have been discovered by Peano, the writer,† and D. R. Curtiss,‡ in which the identical vanishing of the Wronskian implies linear dependence. The starting point for all these investigations is the following:

Theorem A.§ *If $u_1, \cdots u_n$ have at each point of (a, b) finite derivatives of the first $n - 1$ orders, and if among the Wronskians of these functions taken $n - 1$ at a time there is at least one which does not vanish anywhere in (a, b), then, whenever $W(u_1, \cdots u_n) \equiv 0$, $u_1, \cdots u_n$ are linearly dependent throughout (a, b).*

For several years past it has been my custom in my lectures to establish first Theorem A and to deduce from this the two results I and II, using for this purpose an intermediate theorem. As I have not seen this theorem

* Throughout this note we assume x real and speak of an interval (a, b). Everything said applies, however, equally to the case in which x is complex and we have to deal with a continuum in the complex plane. In this latter case, however, I is the *only* result worth recording.

† Trans. Amer. Math. Soc., vol. 2 (1901), p. 139, where references to earlier work by Peano will be found.

‡ Math. Ann., vol. 65 (1908), p. 282.

§ For a simple proof of this theorem see page 140 of my paper just cited.

in the literature, and as I have recently found it useful for its own sake (not merely as a step towards I and II), I venture to give it here in spite of its rather obvious character.

THEOREM B. *If $u_1, \cdots u_n$ have finite derivatives of the first $n-1$ orders at every point of (a, b), then $W(u_1, \cdots u_n) \equiv 0$ is a necessary and sufficient condition that $u_1, \cdots u_n$ be linearly dependent throughout some sub-interval* of (a, b).*

This is obviously (though trivially) true for a single function. We therefore assume that it is true for $n-1$ functions, and use mathematical induction.

Suppose $W(u_1, \cdots u_n) \equiv 0$. Then either

(a) $W(u_1, \cdots u_{n-1}) \equiv 0$, in which case $u_1, \cdots u_{n-1}$ are linearly dependent throughout some sub-interval, and hence the same is true of $u_1, \cdots u_n$; or

(b) there is a point of (a, b) where $W(u_1, \cdots u_{n-1}) \neq 0$, and hence, on account of the continuity of the derivatives of the first $n-2$ orders of the u's, there is a neighborhood of this point at no point of which does $W(u_1, \cdots u_{n-1})$ vanish. Hence in this sub-interval $u_1, \cdots u_n$ are linearly dependent by A. Thus B is proved.

The deduction of I and II from B is as follows:

If $W(u_1, \cdots u_n) \equiv 0$, we know from B that there exist constants $c_1, \cdots c_n$, not all zero, such that throughout a certain sub-interval (a', b')

$$c_1 u_1 + \cdots + c_n u_n \equiv 0. \tag{2}$$

If $u_1, \cdots u_n$ are analytic throughout (a, b), the same is true of the first member of (2). Hence, from the fact that this function vanishes throughout (a', b'), we infer that it vanishes throughout (a, b). Thus I is established.

On the other hand, if $u_1, \cdots u_n$ satisfy (1) at every point of (a, b), the same is true of the first member of (2), which we will call u. But, from (2), we see that u and its derivatives of all orders vanish at any interior point of (a', b'). This, by a fundamental theorem concerning equations of the form (1), implies that u is identically zero throughout (a, b). Thus II is proved.

---

* And hence throughout some sub-interval of every sub-interval of (a, b).

HARVARD UNIVERSITY,
January, 1916.

NOTE ON A THEOREM ON ENVELOPES.

By W. R. Longley.

If a one-parameter family of plane curves $F(x, y, \alpha) = 0$ has an envelope, its rectangular equation is obtained by eliminating α from the two equations $F = 0$ and $F_\alpha = 0$.* Or a parametric representation is obtained by solving the two equations for x and y in terms of α. In order to establish conditions under which this solution is possible appeal must be made to the theory of implicit functions. In a paper in these *Annals* (second series, vol. 12, Jan., 1911, pp. 73–102) Risley and MacDonald have applied standard theorems on implicit functions to determine conditions for the existence of an envelope. In the first part of the discussion it is assumed that the curves are given in the explicit form $y = f(x, \alpha)$, and the authors have stated (p. 86) a " fundamental theorem " which is said to " summarize the facts concerning envelopes of the family of curves given in the explicit form $y = f(x, \alpha)$." This theorem is the following:

"Given a one-parameter family of curves $y = f(x, \alpha)$, where $f(x, \alpha)$ is an analytic function of x and α in the neighborhood of (x_0, α_0), such that

$$(1) \qquad f_{\alpha,\, x^i}(x_0, \alpha) \equiv 0, \qquad i = 0, 1, 2, \cdots, m - 1,$$

but
$$f_{\alpha,\, x^m}(x_0, \alpha) \not\equiv 0,$$

$$(2) \qquad f_{\alpha^j}(x, \alpha_0) \equiv 0, \qquad j = 1, 2, \cdots, n,$$

but
$$f_{\alpha^{n+1}}(x, \alpha_0) \not\equiv 0;$$

then, if

$$(3) \qquad f_{\alpha^{n+1},\, x^m}(x_0, \alpha_0) \neq 0,$$

the curves $y = f(x, \alpha)$ have no envelope in the neighborhood of (x_0, y_0), except a point envelope which occurs whenever $m \geqq 1$; but, if

$$(3') \qquad f_{\alpha^{n+1},\, x^m}(x_0, \alpha_0) = 0,$$

the family $y = f(x, \alpha)$ has an envelope composed of one or more curves through the point (x_0, y_0); also, whenever $m \geqq 1$, a point envelope at that point.

---

* F_α indicates the derivative of F with respect to α. This notation is used throughout the paper.

"When $m = 0$, the notation $i = 0, 1, 2, \cdots, m - 1$ is meaningless and (1) must be replaced by $f_a(x_0, \alpha) \not\equiv 0$."

The form of statement is designed particularly to cover the case when $y_0 \equiv f(x_0, \alpha)$ is an identity in α. Otherwise, as explained in the last sentence of the theorem, the notation of condition (1) has no meaning. When $m \geqq 1$, the theorem as stated is not established by the argument. The hypothesis places certain conditions on the character of the function f in the neighborhood of the "point" (x_0, α_0). The theorem states a conclusion for the neighborhood of the point $P_0(x_0, y_0)$. When $m \geqq 1$, the "point" (x_0, α_0) of the function f is not determined by the geometric point (x_0, y_0). There is no mention of this fact in the proof.

Since it is the "point" (x_0, α_0) that is used throughout the argument, the conclusion applies not to the point (x_0, y_0) but to the curve $y = f(x, \alpha_0)$ at the point (x_0, y_0). The content of the theorem actually proved is: If the hypotheses (1), (2), and (3) are satisfied, the curve $y = f(x, \alpha_0)$ does not touch a branch of the envelope at (x_0, y_0) and no curve of the system for values of α sufficiently near to α_0 touches a branch of the envelope in the neighborhood of (x_0, y_0), except a point envelope which occurs whenever $m \geqq 1$; but if (3′) is satisfied, the family $y = f(x, \alpha)$ has an envelope composed of one or more branches which are tangent at the point (x_0, y_0) to the curve $y = f(x, \alpha_0)$; also, whenever $m \geqq 1$, a point envelope at that point.

If we attempt to obtain the facts for the neighborhood of P_0 the difficulties encountered are considerable. For the ordinary geometric interpretation suppose we consider only real values of x, y, and α. In order to draw the negative conclusion (3) of the theorem quoted it would be necessary to have the hypotheses fulfilled for $x = x_0$ and every real* value of α. It is easy to show by simple examples that if there is one real value of α for which the hypotheses are not satisfied, it is possible to have an envelope which comes within every arbitrarily small neighborhood of P_0. For this purpose consider the system of curves

$$y = f(x, \alpha) = 2\alpha x^2 - \alpha^2 x^3,$$

and inquire about the nature of the envelope near the origin. Here $m = 2$, $n = 0$ and the hypotheses of the theorem are satisfied for every real finite value of α. To examine the value $\alpha = \infty$ we write $\alpha = 1/\beta$ and set $\beta = 0$. Then the hypotheses are not satisfied because the function f is not analytic. The envelope of the system (in addition to the point envelope at the origin) is the straight line $y = x$, which is given in

* This includes $\alpha = \infty$, which would be covered by examining the curve $y = f(x, 1/\beta)$ for $\beta = 0$.

parametric form by the equations $x = 1/\alpha$, $y = 1/\alpha$. For every finite value $\alpha = \alpha_0$ the curve $y = 2\alpha_0 x^2 - \alpha_0^2 x^3$ is tangent to the envelope at the point $x = 1/\alpha_0$, $y = 1/\alpha_0$. By taking α_0 large enough we see that the envelope comes within every arbitrarily small neighborhood of the origin.

A particularly interesting example in this connection is the following:

$$y = \alpha \sin \frac{x}{\alpha}.$$

Every curve of the system passes through the origin, and at this point the hypotheses of the theorem ($m = 1$, $n = 0$) are satisfied for every real value of α except $\alpha = 0$, $\alpha = \infty$. Every curve of the system is tangent at the origin to the line $y = x$. The envelope consists of the infinite number of straight lines given in parametric form by the equations

$$x = \alpha k_i, \qquad y = \alpha \sin k_i, \qquad\qquad i = 1, 2, \cdots,$$

where k_i denote the roots of the equation

$$k_i = \tan k_i.$$

Every curve of the system touches each branch of the envelope twice (once for x positive and once for x negative) and, with the exception of the origin, every branch of the envelope is touched at every one of its points by one curve of the system. In this example we have in every arbitrarily small neighborhood of the origin an infinite number of branches of the envelope, not one of which can be detected by the theorem.

The known theorems on implicit functions demand that the functions and certain derivatives shall be finite for the particular values of the arguments involved. Hence the examples given above are sufficient to show that it is impossible to derive exhaustive criteria for the nature of the envelope in the neighborhood of a point by using the implicit function theorems and the values of f and its derivatives at (x_0, y_0).

SHEFFIELD SCIENTIFIC SCHOOL,
NEW HAVEN, CONN.

NON-ESSENTIAL SINGULARITIES OF FUNCTIONS OF SEVERAL COMPLEX VARIABLES.

By Dunham Jackson.

It is a familiar property of functions of a single complex variable that if $f(z)$ has a pole at the point a, it can be expressed throughout the neighborhood of this point in the form $f(z) = \varphi(z)/(z - a)^m$, where m is a suitable positive integer and $\varphi(z)$ is a function of z which is analytic at a and does not vanish there. Osgood has recently extended this theorem, with appropriate modifications, to functions of several variables. An extension with somewhat different hypotheses, obtained simultaneously and independently by the present writer, is given below. It is assumed that $f(x_1, x_2, \cdots, x_n)$ is analytic throughout the neighborhood of a point, which for convenience is taken as the origin, except for singularities which satisfy a certain requirement of continuity in their distribution,* and which are of such a nature that f is analytic except for poles when regarded as a function of x_1 alone, and it is shown that under these conditions, f can be expressed as the quotient of two functions, each analytic in all the variables in the neighborhood of the origin. The denominator is of course no longer merely a power of a linear factor, in general, and the numerator may vanish at points of the neighborhood in question. The method used is essentially the same as one which is employed in well-known presentations of the proof of Weierstrass's theorem of factorization. The results are expressed in three theorems, the scope of each of which is indicated by the corresponding section-heading. No attempt has been made to obtain the greatest possible generality in the hypotheses, and further generalization is undoubtedly practicable; the aim has been to bring out clearly the central idea of the argument, without unnecessary complications.

1. Non-essential singularity of the first kind. Let the function $f(x_1, x_2, \cdots, x_n)$ be so defined in a region containing the point $(a_1, a_2, \cdots, a_n)$, that if fixed values are assigned to the last $n - 1$ variables, the resulting function of x_1 is analytic except for poles.† The distribution of

* In Osgood's theorem, the singularities are supposed to lie on analytic manifolds.

† Wherever the word *pole* is used in this paper, it will be understood to refer to the behavior of the function under consideration as a function of a single complex variable, when particular values are assigned to such other variables as may be involved. For example, the statement

the poles will be said to be *continuous* at the point $(a_1, a_2, \cdots, a_n)$, if the following condition is fulfilled: There exists a positive number r_1, as small as may be desired, with a corresponding positive number r_2, such that when the inequalities $|x_i - a_i| \leqq r_2$, $i = 2, \cdots, n$, are satisfied, there are no poles on the boundary of the circle $|x_1 - a_1| \leqq r_1$, and the sum of the orders of the poles inside the circle is constant. It will be observed that if this condition is satisfied, and if $f(x_1, a_2, \cdots, a_n)$ has no pole for $x_1 = a_1$, there is no pole at all in the neighborhood of $(a_1, \cdots, a_n)$. Furthermore, an application of the Heine-Borel theorem shows that if E is a closed set of points in the x_1-plane in which $f(x_1, a_2, \cdots, a_n)$ has no pole, and the distribution of poles is continuous at every point of E when $x_2 = a_2, \cdots, x_n = a_n$, then there exists a positive number ρ such that $f(x_1, x_2, \cdots, x_n)$ has no pole in E when $|x_i - a_i| \leqq \rho$, $i = 2, 3, \cdots, n$.

The point about which the discussion centers will be taken for convenience as the point $(0, 0, \cdots, 0)$.

THEOREM I. *Let $f(x_1, x_2, \cdots, x_n)$ be a function of the n complex variables $x_1, x_2, \cdots, x_n$, having the following properties in the neighborhood of the origin:*

(a) If $x_2, \cdots, x_n$, are held fast, $f(x_1, x_2, \cdots, x_n)$ is an analytic function of x_1, except for poles.

(b) The distribution of these poles is continuous at the origin.

(c) Except at the points corresponding to poles, f is analytic in all n variables together.

Then f can be represented throughout a sufficiently restricted neighborhood of the origin, except at the singular points specified above, in the form

$$f(x_1, x_2, \cdots, x_n) = \frac{\varphi(x_1, x_2, \cdots, x_n)}{\psi(x_1, x_2, \cdots, x_n)},$$

where φ and ψ are analytic in all n variables throughout the neighborhood in question, and φ does not vanish there.

The denominator, here and in the later theorems, may be taken as a polynomial in the first variable with coefficients which are analytic functions of the remaining variables.

It is obvious that there is nothing to be proved unless $f(x_1, 0, \cdots, 0)$ has a pole at the point $x_1 = 0$. Let it be assumed, then, that such a pole is present, and that its order is m. If r_1 is a sufficiently small positive quantity, there will be no other pole of the function $f(x_1, 0, \cdots, 0)$, and

that $f(x_1, x_2, \cdots, x_n)$ has no poles in a certain region of the space of the n variables will mean that if $(a_1, a_2, \cdots, a_n)$ is any point of the region, the function $f(x_1, a_2, \cdots, a_n)$ has no pole for $x_1 = a_1$. A special *name* for a singular point of the kind indicated by the title of this section, at which the function becomes infinite with regard to all its arguments together, will not be needed again.

no zero of this function at all, inside or on the circumference $|x_1| = r_1$. In consequence of the hypothesis (b), it is possible to associate with a number r_1 satisfying this condition, a second positive number r_2, so that if

$$(1) \qquad |x_i| \leqq r_2, \qquad i = 2, \cdots, n,$$

the function $f(x_1, x_2, \cdots, x_n)$, considered as to its dependence on x_1 alone, has no pole for $|x_1| = r_1$, and has just m poles for $|x_1| < r_1$, when multiplicities are taken into account; and it may be assumed further that there are no zeros on the circumference $|x_1| = r_1$. For the last statement is surely correct if r_2 is sufficiently small, since f is continuous in all n variables and different from zero for $|x_1| = r_1$, $x_2 = \cdots = x_n = 0$; and a value of r_2 once chosen to satisfy the earlier requirements may be replaced by any smaller value.

It appears that the function f_{x_1}/f is analytic in all n variables for $|x_1| = r_1$, if the last $n - 1$ variables are subject to the restriction (1), as will be assumed henceforth. The integral of this quotient, extended around the circumference, is an analytic function of $x_2, \cdots, x_n$, and retains this character if the integrand is replaced by $x_1^k f_{x_1}/f$, where k is any positive integer. Let us consider first the integral of f_{x_1}/f itself. Except for a factor $2\pi i$, its value is the difference between the number of poles and the number of zeros of f inside the circle. As this difference is an integer, the statement that it is an analytic, and hence continuous, function of $x_2, \cdots, x_n$, is equivalent to the statement that it is constant. But the number of poles has been assumed to be constant, by itself, and hence the number of roots is constant also, and equal to zero, the value which it has for $x_2 = \cdots = x_n = 0$. The only singularities of f_{x_1}/f inside the circle are those due to the poles of f.

The integral of $x_1 f_{x_1}/f$, multiplied by a constant factor, gives the sum of the values of x_1 for which poles occur, each counted according to its multiplicity, or, as may be said for brevity, the sum of the poles. It is an analytic function of $x_2, \cdots, x_n$, as already stated. By integrating $x_1^2 f_{x_1}/f$, it is shown that the sum of the squares of the poles is analytic, and similarly for the higher powers. It follows that if $p_1(x_2, \cdots, x_n)$, $\cdots$, $p_m(x_2, \cdots, x_n)$, are the elementary symmetric functions of the poles, they are analytic* for all values of $x_2, \cdots, x_n$, satisfying (1), and the polynomial of the mth degree in x_1,

$$\psi(x_1, x_2, \cdots, x_n) = x_1^m - p_1 x_1^{m-1} + \cdots \pm p_m,$$

whose roots correspond in location and multiplicity to the poles of f

* Cf. Osgood, Madison Colloquium, 1914, pp. 181–184; Picard, Traité d'analyse, vol. 2, chapter IX, section 8; Goursat, Cours d'analyse mathématique, vol. 2, section 356.

in the neighborhood $|x_1| \leqq r_1$ of the origin, is analytic in all n of its arguments.

The function

$$\varphi(x_1, x_2, \cdots, x_n) = f(x_1, x_2, \cdots, x_n)\psi(x_1, x_2, \cdots, x_n),$$

if suitably defined at the points where it has removable discontinuities as a function of x_1, is analytic in x_1 throughout the circle, for each set of values of $x_2, \cdots, x_n$. But if x_1 is any interior point of the circle, φ is expressible in the form

$$\varphi(x_1, x_2, \cdots, x_n) = \frac{1}{2\pi i}\int_C \frac{\varphi(t, x_2, \cdots, x_n)}{t - x_1}\,dt,$$

where C denotes the circumference. The integrand here is an analytic function of the $n + 1$ variables $t, x_1, x_2, \cdots, x_n$, together, all along the path of integration, and the integral consequently is an analytic function of the last n of these variables. That is, φ is analytic throughout the neighborhood of the origin, and the relation

$$f = \frac{\varphi}{\psi}$$

gives the desired representation of f. It follows immediately from the definition of φ that it does not vanish in the neighborhood of the origin, since the zeros of ψ are cancelled by the poles of f, and zeros of f do not exist.

Considered by itself, the present theorem can be proved more simply. For as soon as it has been established that f is different from zero throughout the neighborhood of the origin, it can be shown at once that the reciprocal of f has only removable singularities. But the reasoning based on the introduction of the function ψ will be of use in the proof of the later theorems.

2. Singularity of the second kind, two variables.

THEOREM II. *Let $f(x, y)$ be a function of the complex variables x, y, having the following properties in the neighborhood of the origin:*

(*a*) *If y is held fast, $f(x, y)$ is an analytic function of x, except at most for poles.*

(*b*) *The distribution of these poles is continuous at every point except the origin itself.*

(*c*) *Except at the origin and at the points corresponding to poles, f is analytic in x and y together.*

Then f can be represented throughout a sufficiently restricted neighborhood of the origin, except at the singular points specified above, in the form

$$f(x, y) = \frac{\varphi(x, y)}{\psi(x, y)},$$

where φ and ψ are analytic in both variables throughout the neighborhood in question.

Let ρ_1 be a positive number such that $f(x, 0)$ is analytic for $|x| = \rho_1$. Then if ρ_2 is sufficiently small, $f(x, y)$ will be analytic for $|x| = \rho_1$, when $|y| \leqq \rho_2$. We shall denote the regions $|x| \leqq \rho_1$, $|y| \leqq \rho_2$, by R_1 and R_2 respectively. Let y_0 be any particular value of y, distinct from zero, in R_2. Suppose that the poles of $f(x, y_0)$ in R_1 are situated at the points x', x'', $\cdots$, $x^{(q)}$. It will be possible to associate with each of the points $x^{(j)}$ a pair of positive numbers $r_1^{(j)}$, $r_2^{(j)}$, in such a way that the circles $|x - x^{(j)}| \leqq r_1^{(j)}$ are wholly exterior to each other and interior to R_1, and that the sum of the orders of the poles in any one circle remains constant as long as $|y - y_0|$ does not exceed the corresponding quantity $r_2^{(j)}$; for the distribution of the poles is assumed to be continuous at each of the points in question. The closed region obtained by subtracting from R_1 the interiors of the small circles, contains no pole of $f(x, y_0)$, and will remain free from poles if y is allowed to vary subject to a suitable inequality $|y - y_0| \leqq r_2^{(0)}$. If r_2 is the smallest of the quantities $r_2^{(0)}$, $r_2^{(j)}$, the sum of the orders of the poles in all R_1 will be constant throughout the neighborhood $|y - y_0| \leqq r_2$. As every point of R_2, except the point $y = 0$, has a neighborhood of this sort, an application of the Heine-Borel theorem leads to the conclusion that the sum of the orders of the poles in R_1 has a single constant value m throughout R_2, except at the one point just specified.

A particular value $y = y_0 \neq 0$ in R_2 being designated once more, the circles about the points $x^{(j)}$ may be supposed chosen so small that none of them contains any zero of $f(x, y_0)$ in its interior or on its boundary. Then the quotient f_{x_1}/f is analytic on the boundaries of these circles for $y = y_0$, and will remain so for $|y - y_0| \leqq r_2$, if r_2 is further diminished sufficiently. The reasoning employed in the proof of the first theorem may be applied to each of the circles in turn; the quotient may be integrated around the circumference, with or without a factor x_1^k. It is found that no zeros come in to complicate the representation of the poles, and that the sum of the poles, or of like powers of the poles, is an analytic function of y. For $|y - y_0| \leqq r_2$, the poles in the small circles are all the poles in R_1. The sum of the kth powers of the poles in R_1, where k is any positive integer, may be obtained by adding the corresponding quantities for the various small circles. The sums of powers, and hence the quantities $p_1(y)$, $\cdots$, $p_m(y)$, the elementary symmetric functions of the poles, are analytic in y for $|y - y_0| \leqq r_2$.

From this it follows that the functions $p_k(y)$ are analytic throughout R_2, except for $y = 0$. But they remain finite even in the neighborhood of this point, for it is apparent from their meaning that their numerical values can not exceed certain readily assignable quantities involving r_1. If they are defined for $y = 0$, without regard to the distribution of the poles of f, by means of their limiting values, they will be analytic in R_2 without exception, and the function

$$\psi(x, y) = x^m - p_1(y)x^{m-1} + \cdots \pm p_m(y)$$

will be analytic in both variables.

The product $\varphi = f\psi$, then, if suitably defined at points of removable discontinuity, is analytic as a function of x throughout R_1, when y has any value different from zero. But it is readily seen that this is true even for $y = 0$, at least if an exception is made of the origin in the x-plane.

The function $\varphi(x, 0)$ is certainly analytic at any point, distinct from the origin, where $f(x, 0)$ has no pole. Let $x_0 \neq 0$, then, be a point where $f(x, 0)$ has a pole of the λth order. It is to be shown that $\psi(x, 0)$ vanishes here λ times. The point x_0 may be surrounded with a circle of arbitrarily small radius, on the circumference of which $\psi(x, 0)$ has no root; and then $\psi(x, y)$ will have no root on the circumference for sufficiently small values of y. By extending the integral of ψ_x/ψ around this circle, it is seen that the number of roots of $\psi(x, y)$ inside is a continuous function of y for $y = 0$. But this is the same as the number of poles of f inside the circle, when $y \neq 0$, and the latter number remains by hypothesis equal to λ when y varies slightly from the value zero. Consequently $\psi(x, 0)$ has just λ roots inside the circle; and as this is true, however small the circle may be, it must be that there is a λ-fold root at x_0.

The function $\varphi(x, y)$ may now be regarded as analytic in x at all points in the neighborhood of the origin, with the single exception of the point $(0, 0)$ itself. That it is analytic in both variables together, except at this point, is shown by means of Cauchy's integral formula, as in the preceding theorem. If $y \neq 0$, the boundary of R_1 may be taken as the path of integration, while if $y = 0$, the path in the case of any particular point x will be a smaller circle surrounding the point and excluding the origin. Being analytic in both x and y except at the isolated point $(0, 0)$, the function φ has at most a removable discontinuity there, and the representation

$$f(x, y) = \frac{\varphi(x, y)}{\psi(x, y)}$$

is of the form desired.

3. Singularity of the second kind, n variables. In the statement of the

next theorem, dealing with the space of the n complex variables x_1, x_2, $\cdots$, x_n, reference will be made to a point-set H, the nature of which it will be convenient to describe in advance. It is situated within a certain neighborhood of the origin, to which the whole discussion is restricted; and its points are finite in number in the space of the variables x_1, x_2, if the values of the remaining $n-2$ coördinates are fixed. To restate the latter requirement at greater length, if the last $n-2$ coördinates of a point of H are to have given values, only a finite number of determinations of x_2 are possible, and corresponding to any one of these there are only a finite number of values of x_1.

THEOREM III. *Let $f(x_1, x_2, \cdots, x_n)$ have the following properties in the neighborhood of the origin:*

(*a*) *If $x_2, \cdots, x_n$, are held fast, f is an analytic function of x_1, except at most for poles.*

(*b*) *The distribution of the poles is continuous, except at the points of a set H, of the sort described above.*

(*c*) *Except at the points of H and those corresponding to poles, f is analytic in all n variables together.*

Then f can be represented throughout a sufficiently restricted neighborhood of the origin, except at the singular points just specified, in the form

$$f = \frac{\varphi}{\psi},$$

where φ and ψ are analytic in all n variables throughout the neighborhood in question.

A new proof is required only if the set H includes the origin, since otherwise Theorem I can be applied at once. In any case, the points where f fails to be analytic in all n variables for $x_2 = \cdots = x_n = 0$, whether poles of $f(x_1, 0, \cdots, 0)$ or points of H, are finite in number, and there will be no difficulty in assigning a circle $|x_1| = \rho_1$, on which f is analytic when the last $n-1$ arguments vanish. It will continue to be analytic on this circle, in all n arguments together, for $|x_i| \leqq \rho_2$, $i = 2, \cdots, n$, if ρ_2 is a sufficiently small positive quantity. Let R_1 denote the region $|x_1| \leqq \rho_1$, R_2 the region in the space of the last $n-1$ variables defined by the inequalities to which they have just been subjected, and H_2 the set of points in R_2 whose coördinates figure as the last $n-1$ coördinates of points of H.

By reasoning entirely similar to that used in proving the preceding theorem, it is shown that the sum of the orders of the poles of f in R_1 is constant in the neighborhood of every point of R_2 which is not a point of H_2. It follows that this number has one and the same value throughout

the whole of R_2, points of H_2 excluded. For if $x_3, \cdots, x_n$ are held fast, it remains constant as x_2 varies, avoiding a finite number of exceptional values; thus it acquires a significance as a function of $x_3, \cdots, x_n$ alone, and it is seen to remain constant for all changes of these variables without exception.

It is shown next that $p_1, \cdots, p_m$, the elementary symmetric functions of the poles of f in R_1, are analytic functions of $x_2, \cdots, x_n$, at all points of R_2 except those which are contained in H_2; the proof given at the corresponding stage in the derivation of Theorem II is applicable here with practically no change. Furthermore, it is obvious from their meaning that these functions remain finite throughout R_2. By abandoning their interpretation with reference to the poles of f, they may be regarded as analytic in x_2, when $x_3, \cdots, x_n$, are held fast, even for the isolated values of x_2 to which points of H_2 correspond. Such a point may be surrounded by a circle in the x_2-plane, on the circumference of which the functions are analytic in all $n - 1$ variables. The argument by means of Cauchy's integral formula, with which the earlier demonstrations were concluded, may be employed here to show that when suitably defined in H_2 the quantities $p_1, \cdots, p_m$, are analytic throughout the whole of R_2. The function

$$\psi = x_1^m - p_1 x_1^{m-1} + \cdots \pm p_m$$

is analytic in $x_1, \cdots, x_n$, throughout the neighborhood of the origin.

Of the product $f\psi$, it can be said at the outset that it is analytic in x_1 throughout R_1, if suitably defined at points of removable discontinuity, for all sets of values of $x_2, \cdots, x_n$ in R_2, except those which belong to H_2. As in the special case with which the second theorem deals, this statement is true even for the exceptional sets of values of $x_2, \cdots, x_n$, if the values of x_1 which give points of H are left out of account. The function $f\psi$ is analytic in all the variables together for all but a finite number of points of R_1, if $x_2, \cdots, x_n$ are given any particular values, and the conditions are suitable for the application of Cauchy's integral formula once more, with the conclusion that $f\psi$ is analytic in all n variables throughout the neighborhood of the origin, except at the points of H. If $x_3, \cdots, x_n$, are held fast, the points of H are isolated in the space of the variables x_1 and x_2; it follows that $f\psi$ is analytic in x_1 (and x_2) at the points of H, if suitably defined there, and a final application of Cauchy's integral formula shows that it is analytic in all the variables. Thus the proof is completed.

Harvard University,
Cambridge, Mass.

A CONGRUENCE OF CIRCLES.*

By Frederick Wahn Beal.

Introduction.

1. This paper treats of a particular type of two-parameter systems of circles whose planes envelope a non-developable surface S. The curvilinear coördinates on S will be u and v and the congruence of circles will be called a congruence C. Any point P of a circle of the congruence is to be a point P of a surface S_1, of a single-parameter family of surfaces. The tangent plane at P to S_1 is to pass through the center of the circle and to make an angle with the plane of the circle, which, at most, is a function of u and v. When this angle is a right angle, the system of circles is cyclic. The existence and some general properties of the congruence C will be proved first. Afterwards the particular case in which the center of the circle is the point at which its plane touches its envelope will be considered and some characteristic properties discovered. In this case the system of circles is called a congruence C_0. The system of circles defining the Backlund transformations of a pseudospherical surface is a congruence of this kind. Congruences C_0 characterize a class of surfaces to which belong also the surfaces S_1 as defined above. We say that each surface S_1 is a transform of S. The Bianchi and Backlund transformations† of pseudospherical surfaces are special cases of the more general transformations defined by congruences C_0.

I. General treatment of congruences C.

2. In order to attack the problem under consideration we will associate with S a trirectangular set of moving axes, in such a way that the z-axis is the normal to S and the xy-plane is the tangent plane to S. The coördinates $(a, b, 0)$ of the center of the circle relative to these axes, the radius R of the circle and the angle ϕ which the normal to S_1 at P makes with the plane of the circle will all, in general, be functions of u and v. The formulæ giving the total displacements of a point P are‡

* Read before the American Mathematical Society, Apr. 24, 1915.

† Cf. Eisenhart, Differential Geometry, (Ginn & Co., Boston, 1909), p. 284. Further references to this work will be given in the form Eisenhart, p. 284.

‡ Eisenhart, p. 166.

$$(1)\quad \begin{cases} \delta x = dx + \xi du + \xi_1 dv + (qdu + q_1 dv)z - (rdu + r_1 dv)y, \\ \delta y = dy + \eta du + \eta_1 dv - (pdu + p_1 dv)z + (rdu + r_1 dv)x, \\ \delta z = dz + (pdu + p_1 dv)y - (qdu + q_1 dv)x, \end{cases}$$

where δx and dx denote total and relative displacements respectively, and ξ, ξ_1, η, η_1, are the *translations* and p, p_1, q, q_1, r, r_1 the *rotations* of the tri-rectangular axes.

The coördinates of a point P on the circle are

$$(a + R\cos\theta,\ b + R\sin\theta,\ 0),$$

and the direction-cosines of the normal to S_1 at P are

$$\cos\phi\sin\theta, \qquad -\cos\phi\cos\theta, \qquad \sin\phi.$$

If the motion of P is to be perpendicular to this direction for all displacements on S in the neighborhood of M, the vertex of the trihedral, the following equation must be satisfied:

$$\cos\phi\sin\theta\,\delta x - \cos\phi\cos\theta\,\delta y + \sin\phi\,\delta z = 0.$$

This equation reduces to

$$\begin{aligned} & R\cos\phi d\theta + [\cos\phi(B\cos\theta - A\sin\theta + Rr) + \sin\phi(qa + qR\cos\theta \\ (2)\qquad & \quad - pb - Rp\sin\theta)]du + [\cos\phi(B_1\cos\theta - A_1\sin\theta + Rr_1) \\ & \quad + \sin\phi(q_1 a + q_1 R\cos\theta - p_1 b - p_1 R\sin\theta)]dv = 0, \end{aligned}$$

where A_i and B_i are defined by the following:

$$A_1 = a_v + \xi_1 - r_1 b, \qquad B_1 = b_v + \eta_1 + r_1 a,$$
$$A = a_u + \xi - rb, \qquad B = b_u + \eta + ra.*$$

By the substitution $t = \tan\theta/2$, this equation becomes a Riccati equation and we have, as is true for cyclic systems, that

THEOREM. *Any four surfaces of the system S_1 meet the circles in four points whose cross-ratio is constant.*

If equation (2) is to possess a solution $\theta(u, v)$, it must satisfy a well-known condition of integrability,† which in this case reduces to the form

$$(3)\qquad T\cos\theta + U\sin\theta + V = 0,$$

where T, U and V are functions of u and v. This equation, if it is not an

* Throughout the paper derivatives with respect to u and v are indicated by subscripts, as a_u and a_v.

† Forsyth, a Treatise on Differential Equations (Macmillan & Co., 1903), p. 282.

identity, possesses at most two distinct solutions which may be real and satisfy equation (2). Hence again as for cyclic systems we have that

THEOREM. *If more than two surfaces S_1 exist, there exist a single infinity and the congruence of circles is a congruence C.*

The necessary and sufficient condition that the system of circles be a congruence C is that θ involve an arbitrary parameter. This means that equation (3) is an identity in θ, and gives the following three equations of condition:

$$\begin{aligned} Rb(pq_1 - p_1q) + (AR_v - A_1R_u)\cos^2\phi - a\sin\phi\cos\phi(qB_1 - q_1B) \\ + b\sin\phi\cos\phi(pB_1 - p_1B) - R^2(p\phi_v - p_1\phi_u) = 0, \end{aligned} \tag{4}$$

$$\begin{aligned} Ra(pq_1 - p_1q) + (B_1R_u - BR_v)\cos^2\phi + a\sin\phi\cos\phi(q_1A - qA_1) \\ + b\sin\phi\cos\phi(pA_1 - p_1A) + R^2(q\phi_v - q_1\phi_u) = 0, \end{aligned} \tag{5}$$

$$\begin{aligned} R^2(pq_1 - p_1q) + (AB_1 - A_1B)\cos^2\phi + a\sin\phi\cos\phi(q_1R_u - qR_v) \\ + b\sin\phi\cos\phi(pR_v - p_1R_u) + Ra(q\phi_v - q_1\phi_u) \\ - Rb(p\phi_v - p_1\phi_u) = 0. \end{aligned} \tag{6}$$

These equations may be readily solved for first order partial derivatives, with respect to u, of three of the four quantities a, b, R, and ϕ, unless ϕ is a right angle, in which case the system is cyclic. Hence, in general, one of the four functions is arbitrary and the other three are given by a system of first order partial differential equations which may be put in the type form. A relation may be assumed between a and b which will cause the center of the circle to generate a particular surface.

3. We will close the discussion of the general congruence C by making a few statements that can be easily verified. They will serve to emphasize the differences between the general case and the particular case of cyclic systems.

None of the following three conditions, all of which are satisfied by cyclic systems, is satisfied by a general system C, but without imposing any restrictions on S, the functions a, b, R, and ϕ can be determined so that anyone of these conditions can be satisfied by a congruence C:

The developables of the congruence of axes of the circles correspond to a conjugate system on S.

The lines joining a point of a circle to the focal points of the congruence of axes of the circles intersect under a right angle. For cyclic systems these lines are the tangents considered in the next statement.

The tangents to S_1 at P that pass through the projections on the tangent plane to S_1 at P of the focal points of the associated congruence of

axis intersect at right angles. This condition is satisfied by congruences C_0.

II. Congruences C_0.

When a and b are identically zero, we will call the envelope of the planes of the circles the surface A. The corresponding congruence of circles has been called a congruence C_0. If the lines of curvature on A are taken as parametric and the x-axis of the trihedral tangent to $v = \text{const.}$, the conditions (4), (5) and (6) reduce to the following:

$$qR_v + \eta_1\phi_u = 0, \tag{7}$$

$$p_1R_u + \xi\phi_v = 0, \tag{8}$$

$$R^2K + \cos^2\phi = 0, \tag{9}$$

where K denotes the total curvature of the surface A.* When the equations obtained from (7) and (8) by eliminating R are solved for the derivatives of ϕ, we have the following:

$$\phi_v = \frac{1}{2}\left[\sec\phi\sqrt{\frac{p_1\eta_1}{q\xi}}\frac{\partial}{\partial u}\log K + \tan\phi\frac{\partial}{\partial v}\log K\right], \tag{10}$$

$$\phi_u = \frac{1}{2}\left[\sec\phi\sqrt{\frac{q\xi}{p_1\eta_1}}\frac{\partial}{\partial v}\log K + \tan\phi\frac{\partial}{\partial u}\log K\right], \tag{11}$$

The condition of integrability of these equations is reducible to

$$\begin{aligned}\frac{\partial}{\partial u}\frac{p_1\eta_1}{q\xi}\frac{\partial}{\partial u}\log K + 2\frac{p_1\eta_1}{q\xi}\frac{\partial^2}{\partial u^2}\log K - \frac{p_1\eta_1}{q\xi}\left[\frac{\partial}{\partial u}\log K\right]^2 \\ = \frac{\partial}{\partial v}\log\frac{q\xi}{p_1\eta_1}\frac{\partial}{\partial v}\log K + 2\frac{\partial^2}{\partial v^2}\log K - \left[\frac{\partial}{\partial v}\log K\right]^2.\end{aligned} \tag{12}$$

This equation characterizes the A surfaces.

That pseudospherical surfaces are A surfaces is evident. To show that other A surfaces exist, we have obtained the following fundamental quantities of a surface of revolution of total curvature $-1/au$ which satisfy equation (12) and the six equations,† equivalent to the Gauss and Codazzi equations of condition:

$$\xi_1 = \eta = r = p = q_1 = 0, \qquad \xi = p_1 = U_1, \qquad \eta_1 = U_2,$$

$$r_1 = \frac{d}{du}\log U_2, \qquad q = \frac{U_2}{au}.$$

* Eisenhart, p. 172.

† Eisenhart, pp. 168, 170.

The quantities U_1 and U_2 are functions of u alone and satisfy the following ordinary differential equations:

$$auU_1U_1' + U_2U_2' = 0,$$

$$auU_1U_2'' - auU_2'U_1' - U_1^3U_2 = 0.$$

5. The coördinates of a point P on the normal to a surface S_1, which we will now call A_1, are

$$(R \cos \theta + t \cos \phi \sin \theta, \qquad R \sin \theta - t \cos \phi \cos \theta, \qquad t \sin \phi),$$

where t is the distance from the surface A_1 to P. The conditions that P be a focal point of this congruence of normals, *i. e.*, that P move tangentially to the normal, are

$$\frac{\delta x}{\cos \phi \sin \theta} = \frac{\delta y}{- \cos \phi \cos \theta} = \frac{\delta z}{\sin \phi},$$

where δx, δy, and δz are the total displacements of P given by (1). These conditions reduce to the equations

$$[RR_u + t\xi \sin \theta \cos \phi + \xi R \cos \theta]du + [RR_v - t\eta_1 \cos \theta \cos \phi + \eta_1 R \sin \theta]dv = 0,$$

$$\begin{aligned}[RR_u \sin \phi \cos \phi \cos \theta &- R\xi \sin \phi \cos \phi \sin^2 \theta \\ &+ t\xi \sin \phi \cos^2 \phi \sin \theta \cos \theta + \xi R \sin \phi \cos \theta - tR\phi_u \cos \phi \sin \theta \\ &+ qR^2 \sin \theta \cos \theta + qtR \cos \phi \sin^2 \theta]du + [RR_v \sin \phi \cos \phi \cos \theta \\ &+ \eta_1 R \sin \phi \cos \phi \sin \theta \cos \theta - t\eta_1 \sin \phi \cos^2 \phi \cos^2 \theta - tR\phi_v \cos \phi \sin \theta \\ &- p_1 R^2 \sin^2 \theta + p_1 tR \cos \phi \sin \theta \cos \theta]dv = 0.\end{aligned}$$

If du and dv are eliminated between these equations, a quadratic in t is obtained for which the coefficient of t^2 and the term not involving t are respectively

$$R \cos^2 \phi \sin \theta[p_1 \sin \theta R_u + q \cos \theta R_v + \sin \theta \cos \theta(p_1\xi + q\eta_1)]$$

and

$$- R^3 \sin \theta[p_1 \sin \theta R_u + q \cos \theta R_v + \sin \theta \cos \theta(p_1\xi + q\eta_1)].$$

Hence the total curvature of A_1, that is the reciprocal of the product of the roots, is $- (\cos^2 \phi)/R^2$. Since this expression does not involve θ it is the same for all the surfaces A_1. Moreover, from (9) it follows that it is also the total curvature of A. Hence we have

THEOREM. *The congruence C_0 defines a transformation of a surface A into a single infinity of surfaces A_1 of the same total curvature as A.*

The surfaces A and A_1 are evidently the focal surfaces of the congruence of lines joining A and A_1. The necessary and sufficient condition that this congruence be normal is that the tangent planes to A and A_1 be perpendicular, *i. e.*, $\phi = 0$.

Under this condition equations (7), (8) and (9) show that A and hence A_1 are pseudospherical surfaces. The surfaces A_1 are Bianchi transforms of the pseudospherical surface A in this case.

6. The linear element of the surfaces A_1 is

$$\begin{aligned} dS^2 = (\xi^2 \cos^2 \theta + R_u^2 + 2\xi \cos \theta R_u + q^2 R^2 \cos^2 \theta \sec^2 \phi) du^2 \\ + 2(R_u R_v + \xi \cos \theta R_v + \eta_1 \sin \theta R_u) du dv \\ + (\eta_1^2 \sin^2 \theta + R_v^2 + 2\eta_1 \sin \theta R_v + p_1^2 R^2 \sin^2 \theta \sec^2 \phi) dv^2. \end{aligned}$$

Not more than two of these surfaces can have an orthogonal parametric system in correspondence with the lines of curvature on surface A unless R is constant. When R is constant, ϕ and K are both constant, and the surfaces A_1 are Backlund transforms of the pseudospherical surface A. Under these conditions the parametric curves on A_1 are lines of curvature. These conditions together with equations (7), (8), (9), (10), and (11) give the following:

THEOREM. *If one of the following four conditions is true, the other three are also true*; ϕ, R, *or* K *constant, lines of curvature on surface* A *correspond to lines of curvature on surfaces* A_1.

III. Transformations of surfaces A.

7. The Bianchi and Backlund transformations of a pseudospherical surface yield pseudospherical surfaces, which are reciprocally related to the original surface in the sense that the latter is one of the Bianchi or Backlund transforms of the transformed surfaces. We will now show that the surfaces A_1 obtained from A under the more general transformation C_0 are so related to A.

We assume that the surface A is subjected to one of these transformations and that the resulting surface is A_1. The lines joining corresponding points on A and A_1 are tangent to a one parameter family of curves on A, which we take for the curves $v =$ const. We choose for the curves $u =$ const. the conjugate system to $v =$ const. If the moving trihedral is chosen so that the x-axis is tangent to the curves $v =$ const. we have*

$$\eta = q_1 = 0, \qquad p\eta_1 = q\xi_1.$$

* Eisenhart, p. 173.

Furthermore, from the choice of $v =$ const., it follows that the function θ for the transformation is equal to zero. Hence from equation (2) we have

$$q \tan \phi + r = 0, \qquad \eta_1 + Rr_1 = 0. \tag{13}$$

Equations (4), (5), and (6) then reduce to the two equations

$$\eta_1 p_1(r^2 + q^2) = \xi q r_1^2, \tag{14}$$

$$q^2\xi(\eta_1 r_{1v} - r_1\eta_{1v}) + \eta_1^2 p_1(rq_u - qr_u) = 0. \tag{15}$$

To compute the fundamental quantities of A_1, we associate with A_1 a moving trihedral whose x-axis is the x-axis of the trihedral for A and has the same direction. The vertices of these trihedrals are M and P. Since A and A_1 are focal surfaces of the congruence of lines made up of the x-axes and since a conjugate system of curves is parametric on A, a conjugate system of curves is also parametric on A_1. If the fundamental quantities of A_1 are indicated by primes, both sides of the first three of the following equations give the total displacements of P in the direction of the axes of the trihedral at P. The last three equations give the total displacements of M in the direction of the axes at M.

$$\xi' du + \xi_1' dv = (R_u + \xi)du + (R_v + \xi_1)dv,$$

$$\eta' du + \eta_1' dv = (Rr \sin \phi - qR \cos \phi)du,$$

$$(Rr \cos \phi + qR \sin \phi)du - \cos \phi(\eta_1 + r_1 R)dv = 0,$$

$$\xi du + \xi_1 dv = (-R_u + \xi')du + (-R_v + \xi_1')dv,$$

$$\eta_1 dv = [(\eta' - Rr') \sin \phi - q'R \cos \phi]du + [-Rr_1' \sin \phi - \cos \phi q_1' R]dv,$$

$$[(\eta' - Rr') \cos \phi + q'R \sin \phi]du + [-Rr_1' \cos \phi + q_1' R \sin \phi]dv = 0.$$

These in virtue of (13) and the equality of the total curvatures of A and A_1 at M and P yield the following:

$$\xi' = R_u + \xi, \qquad \eta' = R(r \sin \phi - q \cos \phi), \qquad \eta_1' = q' = 0,$$

$$\xi_1' = R_v + \xi_1, \qquad r' = r \sin \phi - q \cos \phi, \qquad r_1' = r_1 \sin \phi,$$

$$p' = \frac{q}{\eta_1}(R_v + \xi_1), \qquad p_1' = \frac{p_1}{\xi}(R_u + \xi), \qquad q_1' = r_1 \cos \phi.$$

In order that a system of circles of radius R, with center at P and lying in the tangent planes of A_1 possess a single infinity of transformed urfaces, one of which is A, whose normals make the angle ϕ with the tangent planes to A_1 and are perpendicular to the radius of the circle, it is necessary and sufficient that equations analogous to (4), (5), and (6) be satis-

fied and that $\theta = \pi$ be a solution of an equation analogous to (2). These conditions, with the exception of the equation analogous to (4), may easily be shown to be satisfied. Equation (4) reduces to

$$(\xi R_v - \xi_1 R_u) \cos^2 \phi - R^2(p'\phi_v - p_1'\phi_u) = 0.$$

A similar condition for the surface A is

$$(\xi R_v - \xi_1 R_u) \cos^2 \phi - R^2(p\phi_v - p_1\phi_u) = 0.$$

From these we have that the desired condition is

$$(p - p')\phi_v + (p_1' - p_1)\phi_u = 0.$$

This may be reduced to either one of the two equivalent equations

$$\phi_v R_v = \phi_u R_u \sin^2 \phi,$$

$$\eta_1 p_1(rq_u - qr_u)[p_1\eta_1(r^2 + q^2) - q\xi r_1^2] = 0.$$

The first of these gives a relation between R and ϕ and their derivatives. The second which is equivalent to the first is satisfied by virtue of equation (14).

From the preceding we are able to state the following results:*

THEOREM. *Every surface A_1, obtained from the surface A by means of the congruence C_0, admits of a like congruence C_0. In each case one of the transforms of A_1 is A.*

THEOREM. *Every congruence C_0 has associated with it a complex of circles which is composed of congruences C_0.*

* The latter part of abstract 19 on page 490 of the Bulletin of the American Mathematical Society for July, 1915, is incorrect.

UNIVERSITY OF PENNSYLVANIA.

A CASE OF ITERATION IN SEVERAL VARIABLES.

By Albert A. Bennett.

The Relation between Difference Equations and Iteration.

1. If $F_1(x)$ be a given function, the equation

$$f(x+1) = F_1[f(x)] \tag{1}$$

is a functional equation for $f(x)$, of the form of a difference equation, of the first order. The general difference equation of the first order is of the form

$$f(x+1) = H[f(x), x],$$

where H is a given function of two variables. The difference equation (1) has in any solution $f(x)$ obviously one arbitrary constant at least, since we may choose for the value of $f(x_0)$, say, y_0; the values of $f(x_0 + n)$, where n is any positive (or negative) integer, will then be determinate. There remain, however, an infinite number of parameters in $f(x)$, since the values of $f(x_0 + n)$ for n any real number between zero and one are still arbitrary. If we have any solution $f(x)$, we shall have

$$f(x+2) = F_1\{F_1[f(x)]\} \equiv F_2[f(x)]$$

and if we define $F_n(x)$ for n, a positive integer, by the relation

$$F_n(x) \equiv F_{n-1}[F_1(x)] \equiv F_1[F_{n-1}(x)],$$

we shall then have for n a positive integer,

$$f(x+n) = F_n[f(x)]. \tag{2}$$

We may use (2) for the purpose of defining $F_n(x)$ for n other than a positive integer, if once $f(x+n)$ be defined for all real values of n, and the same feature of arbitrary definition for $0 < n < 1$ appears in $F_n(x)$.

Conversely, starting with $F_1(x)$, we may determine $F_n(x)$ for n a positive or negative integer and define it arbitrarily for every value of n between zero and one. The real solutions of

$$f(x+n) = F_n[f(x)],$$

if $F_1(x)$ be a real function defined for all real values of x, will contain but a single arbitrary constant. The problem of defining or determining $F_n(x)$ from $F_1(x)$ is said to be a problem of iteration. Thus, for the case

of a single variable, the study of iteration is essentially the study of a special type of difference equation of the first order.

We shall find that in the case of several variables, by a mere increase in the number of variables, we may make any system of difference equations each of finite order, equivalent to a problem in iteration. Let us consider a set of p difference equations of orders $q_1, q_2, \cdots, q_p$, respectively. Let one of these equations be of the form

$$f(x+q) = H[f(x+q-1),\ f(x+q-2),\ \cdots,\ f(x+1),\ f(x);\ x].$$

By a change of notation, we may put

$$\begin{aligned} x &\equiv f_{(0)}(x), \\ f(x) &\equiv f_{(1)}(x), \\ f(x+1) &\equiv f_{(2)}(x), \\ &\cdots\cdots \\ f(x+q-1) &\equiv f_{(q)}(x), \end{aligned}$$

so that we have the system

$$\begin{aligned} f_{(0)}(x+1) &= 1 + f_0(x), \\ f_{(1)}(x+1) &= f_2(x), \\ f_{(2)}(x+1) &= f_3(x), \\ &\cdots\cdots\cdots\cdots \\ f_{(q-1)}(x+1) &= f_q(x), \\ f_{(q)}(x+1) &= H[f_{(q)}(x),\ f_{(q-1)}(x),\ \cdots,\ f_2(x),\ f_1(x);\ f_0(x)]. \end{aligned}$$

In this system, we note that x itself no longer appears explicitly. By repeating this process for each of the p equations, and by using throughout a single notation for x, viz., $f_{(0)}(x)$, we obtain a system of $1 + q_1 + q_2 + \cdots + q_p$ difference equations in the same number of unknown functions f, and each of the first order. Thus far, we have implied by our notation, that the number of variables occurring in any of the functions f is one, and that these variables consist simply of x itself. If there be k variables, $x_1, x_2, \cdots, x_k$, appearing in the unknown functions f, and in the functions H, we shall proceed as above and write

$$x_i \equiv f_{(0,\ i)}(x_1, x_2, \cdots, x_k), \qquad i = 1, 2, \cdots, k,$$

and so obtain a system of $k + q_1 + q_2 + \cdots + q_p = m$ difference equations of the first order, in which the x's occur only within the functions f. After this normalization, there will be m equations in m unknown functions, each equation being of the first order, while the unknown functions will

contain k variables, $x_1, x_2, \cdots x_k$. We may next suppose that $k = m$. For if $k < m$, we may introduce as additional parameters into our unknown functions, $m - k$ new variables, $x_{k+1}, x_{k+2}, \cdots, x_m$; while if $m < k$, we introduce new equations of the first order in new unknown functions of the same k variables, and so obtain a system for which the number of functions and the number of equations are equal to the number of variables. The reduction may be performed with respect to one variable at a time. Any system of difference equations of arbitrary orders in which the dependent variables for the greatest values of the arguments appear explicitly, together with certain new parameters, and functions, which may be taken in some simple form, may be reduced to a system of the following type:

$$
\begin{aligned}
& f_{(1)}(x_1 + 1, x_2 + 1, \cdots, x_m + 1) \\
& \qquad \equiv F_{(1)1}[f_{(1)}(x_1, \cdots, x_m), f_{(2)}(x_1, \cdots, x_m), \cdots, f_{(m)}(x_1, \cdots, x_m)], \\
& f_{(2)}(x_1 + 1, x_2 + 1, \cdots, x_m + 1) \\
& \qquad \equiv F_{(2)1}[f_{(1)}(x_1, \cdots, x_m), f_{(2)}(x_1, \cdots, x_m), \cdots, f_{(m)}(x_1, \cdots, x_m)], \\
& \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
& f_{(m)}(x_1 + 1, x_2 + 1, \cdots, x_m + 1) \\
& \qquad \equiv F_{(m)1}[f_{(1)}(x_1, \cdots, x_m), f_{(2)}(x_1, \cdots, x_m), \cdots, f_{(m)}(x_1, \cdots, x_m)],
\end{aligned} \tag{3}
$$

the identity being in terms of the variables $x_1, x_2, \cdots, x_m$. The equations (3) we shall write symbolically as

$$f(x + 1) = F_1[f(x)]. \tag{4}$$

As in the case of one variable, we may regard the equation, now symbolic,

$$f(x + n) = F_n[f(x)],$$

either as defining F_n in terms of $f(x + n)$, or $f(x + n)$ in terms of F_n, the first being a problem of differences, and the second, of iteration.

A Certain Class of Series and their Iteration.

2. We shall investigate equations (3) in the case in which $F_{(i)1}[u_1, u_2, \cdots, u_m]$ is of the form

$$F_{(i)1}[u_1, u_2, \cdots, u_m] \equiv au_i + ((u^2)), \qquad i = 1, 2, \cdots, m, \tag{5}$$

where by $((u^k))$ is meant a power series in the m variables $(u_1, u_2, \cdots, u_m)$, for which no term appears of order less than k in the m variables taken together. In (5), the constant a is supposed to be the same for each value of $i = 1, 2, \cdots, m$, and for the present we shall suppose that $|a| \neq 0, \neq 1$.

Let us suppose that we are given an arbitrary constant γ, such that $|\gamma| \neq 0$, and let us seek to determine the coefficients of m series, C, where $C_{(i)}$ is of the form

$$C_{(i)}[u_1, u_2, \cdots, u_m] \equiv \gamma u_i + ((u^2)), \quad i = 1, 2, \cdots, m,$$

such that the C's satisfy the symbolic equation

$$C[F_1(u)] = F_1[C(u)].$$

The coefficients of C will be completely determined in terms of γ, and of the coefficients of F_1. For let us form a product of h u's distinct or different and call this product U. Then U will occur once and only once in each $F_{(i)1}$ with coefficients, respectively, say, d_i, and once and only once in each $C_{(i)}$ with coefficients, say, c_i. The coefficient of the term in U in $C_{(i)}[F_1(u)]$ will be of the form

$$\gamma d_i + c_i a^h + (\textit{terms containing neither } d_i \textit{ nor } c_i).$$

The coefficient of the term in U in $F_{(i)1}[C(u)]$ will be of the form

$$a c_i + d_i \gamma^h + (\textit{terms containing neither } d_i \textit{ nor } c_i).$$

Furthermore c_i occurs for the first time in either $C[F_1(u)]$ or $F[C(u)]$ with the term in U, while all of the other terms in the coefficient of U, except that one containing d_i will contain exclusively coefficients of C and F, which appeared in these respective sets of series with terms in u's of degree less than h. Since we are supposing that the coefficients of F are known, and that, for the induction, the coefficients in C of terms in u's of degree less than h, are already determined, we have, to determine c_i, an equation of the form

$$c_i(a^h - a) = d_i(\gamma^h - \gamma) + (\textit{known terms}).$$

Since $|a| \neq 0, \neq 1$, we see that c_i, also, is completely determined. The above argument does not apply to the initial term, and it is readily verified that γ is indeed arbitrary provided that $|\gamma| \neq 0$.

Now $F_n(u)$ must be, itself, a special case of $C(u)$, since

$$F_n[F_1(u)] = F_1[F_n(u)] = F_{n+1}(u).$$

For n an integer, the initial terms of $F_n(x)$ are such that

$$F_{(i)n} = a^n u_i + ((u^2)), \qquad i = 1, 2, \cdots, m. \tag{6}$$

We may choose for n other than an integer, say,

$$F_{(i)n} = a^{n+P(n)} u_i + ((u^2)),$$

where $P(n)$ is any periodic function of n of period unity and vanishing for $n = 0$. For the sake of simplicity, we shall choose for our definition of $F_{(i)n}$, a series of the form (6) for all values of n, a choice which will be seen to be possible. In other words we shall choose for $F_n(u)$ the series $C(u)$ for which $\gamma = a^n$. We shall need to use the fact that the form of $F_{(i)n}(u)$ is completely determined by this choice, and that the coefficient of any term in $F_{(i)n}(u)$ is a polynomial in a^n, while n enters in no other manner into the series.

Application of Newton's Interpolation Formula.

3. We shall now apply Newton's interpolation formula to the expansion of $F_{(i)n}(u)$, and we shall prove that the series obtained from this formula, in a certain manner, about to be described, is not only, itself, a power series, but is identically equal to the form which would be obtained by the method of undetermined coefficients, as just considered.

Newton's interpolation formula, for $y = f(x)$, when the values $y_i = f(x_i)$ are known for $i = 0, 1, 2, 3, \cdots$, is the following:

$$
\begin{aligned}
(7) \qquad f(x) = y_0 &+ (x - x_0)[x_0x_1] + (x - x_0)(x - x_1)[x_0x_1x_2] \\
&+ (x - x_0)(x - x_1)(x - x_2)[x_0x_1x_2x_3] + \cdots,
\end{aligned}
$$

where

$$
[x_0x_1] = \frac{y_1 - y_0}{x_1 - x_0}, \qquad [x_1x_2] = \frac{y_2 - y_1}{x_2 - x_1}, \qquad [x_2x_3] = \frac{y_3 - y_2}{x_3 - x_2},
$$

$$
[x_0x_1x_2] = \frac{[x_1x_2] - [x_0x_1]}{x_2 - x_0}, \qquad [x_1x_2x_3] = \frac{[x_2x_3] - [x_1x_2]}{x_3 - x_1},
$$

$$
[x_0x_1x_2x_3] = \frac{[x_1x_2x_3] - [x_0x_1x_2]}{x_3 - x_0}.
$$

The only fact that we shall use concerning this formula is that it reduces identically to y_n if we put $x = x_n$. To apply (7) to the question in hand, we shall put $y = f(x) \equiv F_{(i)n}(u)/a^n$, $x = a^n$, for every value of n, while the values of x_j, $j = 0, 1, \cdots$, will be, respectively, $a^0, a^1, \cdots, a^j, \cdots$. The formula then becomes,

$$
\begin{aligned}
(8) \qquad \frac{F_{(i)n}}{a^n} = \Bigg\{ &\frac{F_{(i)0}}{1} + \frac{(a^n - 1)}{(a - 1)}\left(\frac{F_{(i)1}}{a} - \frac{F_{(i)0}}{1}\right) \\
&+ \frac{(a^n - 1)(a^n - a)}{(a - 1)(a^2 - 1)}\left(\frac{F_{(i)2}}{a^3} - \frac{F_{(i)1}}{a^2} - \frac{F_{(i)1}}{a} + \frac{F_{(i)0}}{1}\right) + \cdots \Bigg\}
\end{aligned}
$$

or as it may be written,

$$F_{(i)n} = a^n\left[V_0 + \frac{a^n - 1}{a - 1} V_1 + \frac{(a^n - 1)(a^n - a)}{(a - 1)(a^2 - 1)} V_2 + \cdots + \frac{(a^n - 1) \cdots (a^n - a^{j-1})}{(a - 1) \cdots (a^j - 1)} V_j + \cdots \right], \tag{9}$$

where $V_0 \equiv F_{(i)0}$ *and*

$$V_j(u_1, u_2, \cdots, u_m) \equiv \frac{V_{j-1}[F_{(1)1}, F_{(2)1}, \cdots, F_{(m)1}]}{a^j} - V_{j-1}(u_1, u_2, \cdots, u_m).$$

We shall first show that $V_j(u_1, u_2, \cdots, u_m)$ is of the form $((u^{j+1}))$. This is obviously true for $j=0$, and we shall assume it proved for $j-1=1$, $2, \cdots, h$. Let us take a term in U of degree h in the u's, and let the coefficient of U in V_h be b. The terms obtained from U by replacing u_i by $F_{(i)1}$, $i = 1, 2, \cdots, m$ in V_h, will, with one exception, be of degree higher than h. There will be a single term obtained from U of degree h of the form $b(a^hU)$. Hence in $V_h[F_{(1)1}, \cdots, F_{(m)1}]/a^{h+1} - V_h(u_1, \cdots, u_m)$, the term in U will disappear, and since the same applies to every term of order h, we see that for $j - 1 = h + 1$, the above theorem as to the order of V_j is true. Hence every term of order h in the u's, which appears in the formula (9) for $F_{(i)n}$, will be obtained from V's of subscript less than h. The coefficient of any term of order h in (9) is therefore a polynomial in a^n and is rational except for a^n in the coefficients of $F_{(1)1}$, $F_{(2)1}$, $\cdots$, $F_{(m)1}$. For n any positive integer (9) reduces to an identity and so must coincide with the series for $F_{(i)n}$ given by the method of undetermined coefficients. Since n enters in (9) only in the form a^n, the polynomial in a^n which is in (9) the coefficient of a term in U will coincide for an infinite number of distinct values of a^n, viz., $a^0, a^1, a^2, \cdots, a^j, \cdots$, with the corresponding polynomial obtained by the method of undetermined coefficients. These polynomials must therefore coincide, and the expressions obtained by the two different methods must be identical.

The Solution of the Difference Equation in Terms of the Iteration Series.

4. Having obtained a form for F_n, we may now secure a set of solutions $f_{(i)}$ of (3). For $|a| < 1$, we shall have a definite set of series as the limit of $a^{-n}F_n$, as n becomes positively infinite, while for $|a| > 1$, the same formal series are obtained as the limit of a^nF_n for $n = +\infty$. The series in each case are

$$G_{(i)}(u_1, u_2, \cdots, u_m) \equiv \left[V_0 - \frac{1}{a - 1} V_1 + \frac{1 \cdot a}{(a - 1)(a^2 - 1)} V_2 - \frac{1 \cdot a \cdot a^2}{(a - 1)(a^2 - 1)(a^3 - 1)} V_3 + \cdots \right], \tag{10}$$

as is seen immediately. The G's satisfy the functional equations

$$aG_{(i)}(u_1, \cdots, u_m) = G_{(i)}[F_{(1)1}, F_{(2)1}, \cdots, F_{(m)1}], \tag{11}$$

since, for example,

$$\lim_{n=+\infty} a(a^{-n}F_n) = \lim_{n=+\infty} a^{-n}F_n(F)$$

and similarly for limit, $n = -\infty$. Hence we have,

$$1 + \log_a G_{(i)}(u) = \log_a G_{(i)}[F_1]$$

or, replacing $\log_a G_{(i)}(u)$ by a new notation, $g_{(i)}(u)$,

$$1 + g_{(i)}(u) = g_{(i)}[F_1]. \tag{12}$$

We shall be obliged to restrict the set $g_{(i)}(u)$ to the extent that we shall suppose the system

$$\begin{aligned} x_1 &= g_{(1)}(u_1, u_2, \cdots, u_m), \\ x_2 &= g_{(2)}(u_1, u_2, \cdots, u_m), \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_m &= g_{(m)}(u_1, u_2, \cdots, u_m), \end{aligned} \tag{13}$$

to possess an inverse, which we shall denote by

$$\begin{aligned} u_1 &= g_{(1)}^{-1}(x_1, x_2, \cdots, x_m), \\ u_2 &= g_{(2)}^{-1}(x_1, x_2, \cdots, x_m), \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ u_m &= g_{(m)}^{-1}(x_1, x_2, \cdots, x_m). \end{aligned} \tag{14}$$

Combining (12), (13), and (14), we have

$$\begin{aligned} &g_{(i)}^{-1}[1 + x_1, 1 + x_2, \cdots, 1 + x_m] \\ &\qquad = F_{(1)1}[g_{(1)}^{-1}(x_1, \cdots, x_m), \cdots, g_{(m)}^{-1}(x_1, \cdots, x_m)]. \end{aligned} \tag{15}$$

Hence, the set $g_{(i)}^{-1}$, $i = 1, 2, \cdots, m$, constitute a solution of (3).

Consideration of Some Limiting Cases.

5. When a becomes equal to unity, the expression (9) assumes an indeterminate form, which may, however, be readily evaluated. A form for $F_{(i)n}$ when $a = 1$, may also be obtained by a direct method analogous to that used above. We shall have, for $a = 1$,

$$\begin{aligned} F_{(i)n} = \Big\{ &F_{(i)0} + \frac{n}{1}(F_{(i)1} - F_{(i)0}) \\ &+ \frac{n(n-1)}{1 \cdot 2}(F_{(i)2} - 2F_{(i)1} + F_{(i)0}) + \cdots \Big\} \end{aligned} \tag{16}$$

or as it may be written

$$(17)\quad F_{(i)n} = \left[V_0 + \binom{n}{1}V_1 + \binom{n}{2}V_2 + \cdots + \binom{n}{j}V_j + \cdots\right],$$

where $V_0 = F_{(i)0}$ *and*

$$V_j = F_{(i)j} - \binom{j}{1}F_{(i)j-1} + \binom{j}{2}F_{(i)j-2} - \cdots \pm F_{(i)0}.$$

We shall not consider the case in which $a = 0$. If $|a| = 1$, but $a^j \neq 1$, $j = 1, 2, \cdots$, our original formulæ hold. If, however, $|a| = 1$, and also $a^j = 1$, for some positive integer j, a new case occurs. Let h be the smallest positive integer (different from zero), for which $a^h = 1$. Then the first h terms of (8) and (9) present no difficulty. The $(h + 1)$st term contains however in the denominator the factor $(a^h - 1)$, while this factor will also occur in the numerator of this term, for n any positive integer. If we break off (8) after the hth term the formula will continue to be an identity for $n = 0, 1, 2, \cdots, h - 1$. If $F_{(i)h} = F_{(i)0}$, in which case we always have $a^h = 1$, then the first h terms of (8) will continue to represent $F_{(i)n}$ for every integer value of n, and will therefore, by the general argument, represent a possible definition of $F_{(i)n}$ for all values of n. If $F_{(i)h} \neq F_{(i)0}$, we cannot define $F_{(i)n}$ for all values of n as a series of the type we are considering.*

Discussion of the Convergence of the Series.

6. We shall now demonstrate that (8) is convergent when $F_{(i)1}$ is, $i = 1, 2, \cdots, m$. Let us suppose first that $|a| > 1$. Let us suppose $F_{(i)1}$ dominated by a function $\tilde{F}_{(i)1}$ all of whose coefficients are real and positive. From the method of obtaining the coefficients of $F_{(i)2}$ from those of $F_{(1)1}, F_{(2)1}, F_{(3)1}, \cdots, F_{(m)1}$, we see that the coefficients of $\tilde{F}_{(i)2}$ will be real and positive and will dominate those of $F_{(i)2}$. Similarly for n a positive integer $F_{(i)n}$ will be dominated by $\tilde{F}_{(i)n}$. Now, if $\tilde{F}_{(i)n}$ be determined by (8) as a convergent series for all values of n within a certain region, we see that since $\tilde{a}^n$ is an increasing function of n for $\tilde{a} > 1$, $\tilde{F}_{(i)n}$ will be an increasing function of n for any given values of the independent variables, and in fact each term of (9) is an increasing function of n. We must show that for any set of functions, $\tilde{F}_{(1)1}, \tilde{F}_{(2)1}, \cdots, \tilde{F}_{(m)1}$, whose expansions in power series contain no term with negative coefficients, the $\tilde{V}$'s, obtained as in (9), will have no negative coefficients. Let us suppose, for induction, that $\tilde{V}_{j-1}$ has no negative terms, then we may write

* Cf. the detailed discussion in the case of one variable, by the author, *Annals of Math.* vol. 17 (1915), p. 35 ff.

$$\tilde{V}_j(u_1, u_2, \cdots, u_m) = \frac{1}{\tilde{a}^j}\{V_{j-1}(\tilde{a}u_1, \cdots, \tilde{a}u_m) + \varphi\} - V_{j-1}(u_1, u_2, \cdots, u_m),$$

where φ consists of partial derivatives of $V_{j-1}(\tilde{a}u_1, \cdots, \tilde{a}u_m)$ multiplied by products of $[F_{(i)1}(u_1, u_2, \cdots, u_m) - \tilde{a}u_i]$ for different or the same values of i, by simply expanding $V_{j-1}(F_{(1)1}, F_{(2)1}, \cdots, F_{(m)1})$ by Taylor's formula for the case of m variables. Now every term of φ has a non-negative coefficient, and since $\tilde{a} > 1$, and $V_{j-1}(u_1, u_2, \cdots, u_m)$ contains, for terms of lowest degree, those of degree j, we see that $(1/a^j)V_{j-1}(\tilde{a}u_1, \tilde{a}u_2, \cdots, \tilde{a}u_m) - V_{j-1}(u_1, u_2, \cdots, u_m)$ contains no terms with negative coefficients. The coefficient in (9) of V_j is itself positive, since, as we have remarked, a^n is an increasing function of n. It therefore suffices to establish the convergence of (9) for a set of dominant functions.

Putting $\tilde{F}_{(i)1} = \tilde{a}u_i/(1 - \tilde{b}s)$, where $s = u_1 + u_2 + \cdots + u_m$, we have

$$\text{(18)} \quad \tilde{V}_j = \frac{\tilde{b}^j \cdot 1 \cdot (1 + \tilde{a})(1 + \tilde{a} + \tilde{a}^2) \cdots (1 + \tilde{a} + \cdots + \tilde{a}^{j-1})s^j u_i}{[1 - \tilde{b}s][1 - \tilde{b}(1 + \tilde{a})s][1 - \tilde{b}(1 + \tilde{a} + \tilde{a}^2)s] \cdots [1 - \tilde{b}(1 + \tilde{a} + \cdots + \tilde{a}^{j-1})s]}.$$

But the series $\tilde{F}_{(i)n}$ may be summed in this case in finite terms. We have, by inspection, a possible $\tilde{F}_{(i)n}$ in

$$\text{(19)} \qquad \tilde{F}_{(i)n} = \frac{\tilde{a}^n u_i}{1 - \tilde{b}\dfrac{1 - \tilde{a}^n}{1 - \tilde{a}}s},$$

which must coincide with the sum of the series given by (9). Since (18) converges when written as a power series, for $\tilde{b}(\tilde{a}^n - 1)|s| < (\tilde{a} - 1)$, the set of series $F_{(i)1}$, dominated by $\tilde{a}u_i/(1 - \tilde{b}s)$, $\tilde{a} > 1$, will converge at least within this region. For the case $|a| < 1$, we need merely invert the series to obtain $a' = 1/a$ such that $|a'| > 1$, while a' takes the place of a in the inverted series. For $a = 1$, the expression (18) continues to hold while (19) becomes

$$F_{(i)n} = \frac{u_i}{1 - n\tilde{b}s},$$

which converges for $|n|\tilde{b}|s| < 1$. The convergence for $|a| = 1$, but $a^h \neq 1$, $h = 1, 2, 3, \cdots$, depends upon other conditions as well as the convergence of $F_{(i)1}$ as is seen in the case $m = 1$, discussed in the previous paper.

Princeton University,
October, 1915.

THE ARITHMETIC GENUS OF AN ALGEBRAIC MANIFOLD IMMERSED IN ANOTHER.

By S. Lefschetz.

Introduction.

1. Let V_r be an r-dimensional irreducible algebraic variety[1] immersed in an $(r+k)$-fold space S_{r+k}, in which it has no singularities, or else is the projection of a non-singular variety in a higher space. There can be found[2] in V_r a fixed number ρ of *hypersurfaces*, $C_1, C_2, \cdots, C_\rho$, such that to any other C, there corresponds a relation of *equivalence* in the sense of Severi, of the type

$$\lambda C = \lambda_1 C_1 + \lambda_2 C_2 \cdots + \lambda_\rho C_\rho,$$

where the λ's are integers, while no such relation exists between $C_1, C_2, \cdots, C_\rho$. These last hypersurfaces form a *base* in V_r, and the λ's may be called the *base characters* of C. The object of this paper is *to find the arithmetic genus of a manifold complete intersection in V_r, in terms of the base characters of the hypersurfaces which determine it.* The case of $r = 1$, $k = 2$ was first derived by Salmon,[3] that of $r = 1$, $k = n - 1$ by Veronese,[4] that of $r = 2$, $k = n - 2$ by Severi.[5] For an algebraic surface the problem has already been solved by Severi,[6] but as his method admits of no easy generalization, the one followed here is completely different, and consists in the extensive use of a symbolism, already somewhat employed by the Italian geometer.[7]

In part I, some auxiliary properties are considered, notably the extension of certain formulas to virtual manifolds. In part II the solution of the fundamental problem is given, while in part III are found applications to loci of flats and to complete intersections in an S_r.

---

[1] A d-flat will be designated throughout by S_d. The name *variety* will be reserved for V_r. Any of its subvarieties will be called a *manifold*. If the manifold has $r - 1$ dimensions, it will be called a *hypersurface*.

[2] *Severi*, La base minima. . . . Ann. Ec. Norm. Sup. (1908), p. 467.

[3] Geometry of three dimensions. 4th edition (1882), p. 309–311.

[4] Über die Methode des Projicirens und Schneidens. Math. Ann., vol. 19 (1882), p. 203.

[5] Su alcune questioni di postulazione. Rendic. del Circ. Mat. di Palermo, vol. 17 (1903), p. 76–79.

[6] Sulla base per la totalita delle curve. . . . Math. Ann. (1906), p. 221.

[7] Fondamenti per la geometria sulle varieta algebriche. Rendic. del Circolo Mat. di Palermo (1909), p. 48, footnote.

The assumption is made throughout that all manifolds considered have no assigned singularities.

2. Let $v(l)$ be the *postulation* of V_r, that is the number of conditions imposed upon a V_{r+k-1}, of order l sufficiently high, when it goes through V_r. According to Hilbert[8]

$$v(l) = k_0\binom{l+r}{r} + k_1\binom{l+r-1}{r-1} + \cdots + k_{r-1}(l+1) + k_r,$$

the k's being certain constants. The *arithmetic genus*,[9] or simply *genus*, of V_r, which will be designated by $[V_r]$, when $r > 0$, is given by

$$[V_r] = (-1)^r(k_0 + k_1 + \cdots + k_r - 1) = (-1)^r(v(0) - 1).$$

When $r = 0$, V_r is a group of points, and the genus $[V_r]$, by definition, the number of points in the group.

The letters $A, B, C, \cdots$ will be used to designate hypersurfaces, and in accordance with the notion of equivalence already referred to, two hypersurfaces belonging to the same continuous algebraic system will be designated by the same letter. Let $A, B, \cdots, C$, be a certain number h of hypersurfaces. If their complete intersection has exactly $r - h$ dimensions, it will be designated by $AB \cdots C$.[10] When no confusion may arise, we will sometime designate it by M_{r-h}, or M, and write $M = AB \cdots C$. What is meant by a hypersurface being the sum of two others is well known, as well as the notation to indicate it. This will be extended here, and we will denote the intersection of M with the sum of B and C by $M(B + C)$. The genus of $M = AB \cdots C$ will be denoted by $[M]$, or by $[AB \cdots C]$.

In addition to these notations, the following will be introduced: Let $F(A, B, \cdots, C)$ be a polynomial, or a series proceeding according to positive integral powers of $A, B, \cdots, C$. Then

(*a*) By $[F(A, B, \cdots, C)]$ we will denote the result obtained when the constant term is left unchanged, and when in place of the term in $A^aB^b \cdots C^c$ we put

$[A^aB^b \cdots C^c]$, if its degree is inferior to r;

zero, if its degree exceeds r.

(*b*) By $\{F(A, B, \cdots, C)\}$ we will designate the same result as the pre-

---

[8] Über die Theorie der algebraischen Formen. Math. Ann., vol. 36 (1890), p. 512, 520.

[9] *Severi*. Fundamenti. . . . p. 41. The definition given here differs from his only when $r = 0$.

[10] In § 5 a meaning will be attributed to $AB \cdots C$, even when this complete intersection is of less than $r - h$ dimensions. In this case, however, there is no connection between the actual intersection and the symbol $AB \cdots C$.

ceding, as far as terms of degree less than r are concerned, but in place of $A^a B^b \cdots C^c$ we will set

$$[A^a B^b \cdots C^c] - 1, \text{ if the degree is } r,$$

$$(-1)^{k-1}, \text{ if the degree is } r + k, \ k > 0.$$

In particular if $M = A^a B^b \cdots C^c$, with $h = a + b + \cdots + c$, then when $h \leqq r$, $[M]$, $\{M\}$, are the genus as defined above, or as defined by Severi, respectively, while when $h > r$, $[M] = 0$, $\{M\} = (-1)^{r-h-1}$.

From the above follows

$$\{F(A, B, \cdots, C)\} = [F(A, B, \cdots, C)] + N,$$

where N is a certain constant. To obtain its value, remark that the term $\alpha \cdot A^a B^b \cdots C^c$ will contribute to N an amount

$$(-1)^{k-1} \cdot \alpha, \text{ if its degree is } r + k, \ k \geqq 0;$$

$$\text{zero, if its degree is less than } r.$$

Hence, if $F(x) = \Sigma \alpha_n x^n$ is the result obtained when in $F(A, B, \cdots, C)$ we make $A = B = \cdots = C = x$, then

$$(-1)^{r-1} N = \sum_{n=r} (-1)^n \alpha_n = F(-1) - \sum_{n=0}^{r-1} (-1)^n \alpha_n.$$

This result will be useful later.

I. Preliminary formulas. Virtual manifolds.

3. Let[11] A, C_1, C_2, be hypersurfaces in V_r, such that $A = C_1 + C_2$. Then whatever r may be

$$\{A\} = \{(1 + C_1)(1 + C_2) - 1\}. \tag{1}$$

For if $r > 1$, this is a formula of Severi's,[12] since $\{M\}$ is the genus of M according to his definition in that case. If $r = 1$, $\{A\} = [A] - 1$, $\{C_1\} = [C_1] - 1$, $\{C_2\} = [C_2] - 1$, $\{C_1 C_2\} = +1$, and the formula reduces to

$$[A] = [C_1 + C_2],$$

which is true, since A, C_1, C_2, are groups of points, and their genus, the number of points in each group.

Let $M = B_1 \cdots B_h$ be a complete intersection in V_r, $h \leqq r - 1$. If we apply (1) in M, we get

$$\{MA\} = \{M((1 + C_1)(1 + C_2) - 1)\}. \tag{2}$$

---

[11] Restrictions upon the formulas of this and the next paragraph, will be considered in § 5.

[12] Fundamenti. . . , p. 42.

If $h = r + k$, $k \geqq 0$, the quantities involved are mere symbols. From § 2 follows, however, $\{MA\} = \{MC_1\} = \{MC_2\} = -\{MC_1C_2\} = (-1)^k$ and the relation is still correct. Hence, if $F(A, B, \cdots, C)$ is a polynomial or a power series

$$\{F(A, B, \cdots, C)\} = \{F((1 + C_1)(1 + C_2) - 1, B, \cdots, C)\}. \quad (3)$$

From this follows easily the generalization of (1). I say that if $A = \sum_{i=1} C_i$, then

$$\{A\} = \left\{\prod_{i=1}^{n} (1 + C_i) - 1\right\}. \quad (4)$$

For, first this is true if $n = 2$. Grant that it holds when A is the sum of less than n hypersurfaces. Then, if $C' = C_{n-1} + C_n$,

$$\{A\} = \left\{(1 + C')\prod_{i=1}^{n-2} (1 + C_i) - 1\right\} = \left\{(1 + C_{n-1})(1 + C_n)\prod_{i=1}^{n-2} (1 + C_i) - 1\right\}$$

which is the formula to be proved.

From (4) follows the generalization of (3), thus: If A designates the same sum as in (4),

$$\{F(A, B, \cdots, C)\} = \left\{F\left(\prod_{i=1}^{n} (1 + C_i) - 1, B, \cdots, C\right)\right\} \quad (5)$$

which is derived from (4), like (3) from (1).

Finally if $A_i = \sum_{k=1}^{n_i} C_{ik}$, $(i = 1, 2, \cdots, h)$, applying (5) we obtain

$$\{A_1A_2 \cdots A_h\} = \left\{\prod_{i=1}^{h}\left(\prod_{k=1}^{n_i} (1 + C_{ik}) - 1\right)\right\}. \quad (6)$$

4. From § 2 and formula (4) follows, that if $A = \sum_{i=1}^{n} C_i$,

$$\{A\} = \left\{\prod_{i=1}^{n} (1 + C_i) - 1\right\} = \left[\prod_{i=1}^{n} (1 + C_i) - 1\right] + N$$

where, in this case, N is a certain integer, and from what has been shown in that paragraph,

$$(-1)^{r-1}N = ((1 + x)^n - 1)_{x=-1} - \sum_{k=1}^{r-1} (-1)^k \binom{n}{k};$$

$$\therefore \qquad N = (-1)^r \sum_{k=0}^{r-1} (-1)^k \binom{n}{k} = -\binom{n-1}{r-1};$$

$$\therefore \qquad \{A\} = \left[\prod_{i=1}^{n} (1 + C_i) - 1\right] - \binom{n-1}{r-1}, \quad (7)$$

formulas which will be of frequent use.[13]

[13] For the simplification involved, see Netto, Lehrbuch der Kombinatorik, p. 248, formula 14.

We have next, from (6), and with the same notations,

$$\{A_1A_2\cdots A_h\} = \Big[\prod_{i=1}^{h}\Big(\prod_{k=1}^{n_i}(1+C_{ik})-1\Big)\Big]+N',$$

where N' is again an integer. This integer remains the same if all the C's coincide with the same hypersurface C. Let M be what $A_1A_2\cdots A_k$ becomes in that case, and $n = \Sigma n_i$. We have

$$\begin{aligned}\{M\} &= \Big\{\prod_{i=1}^{h}[(1+C)^{n_i}-1]\Big\}\\ &= \{(1+C)^n - \sum_i (1+C)^{n-n_i} + \sum_{i,j,\,i\neq j}(1+C)^{n-n_i-n_j}+\cdots+(-1)^h\}.\end{aligned}$$

From (4) and (7) follows

$$1+\{\overline{kC}\} = \{(1+C)^k\} = [(1+C)^k] - \binom{k-1}{r-1};$$

$$\begin{aligned}\therefore\quad \{M\} = {} & [(1+C)^n - \sum_i (1+C)^{n-n_i} + \sum_{i,j,\,i\neq j}(1+C)^{n-n_i-n_j}+\cdots+(-1)^h]\\ & -\Big(\binom{n-1}{r-1} - \sum_i\binom{n-n_i-1}{r-1} + \sum_{i,j,\,i\neq j}\binom{n-n_i-n_j-1}{r-1}\cdots\\ & + (-1)^{h-1}\sum_i\binom{n_i-1}{r-1}\Big).\end{aligned}$$

The last parenthesis is equal to $-N'$.

$$\begin{aligned}\therefore\quad \{A_1A_2\cdots A_h\} = {} & \Big\{\prod_{i=1}^{h}\Big(\prod_{k=1}^{n_i}(1+C_{ik})-1\Big)\Big\} - \binom{n-1}{r-1}\\ & + \sum_i\binom{n-n_i-1}{r-1} - \sum_{i,j,\,i\neq j}\binom{n-n_i-n_j-1}{r-1} \qquad (8)\\ & + \cdots + (-1)^h\sum_i\binom{n_i-1}{r-1}.\end{aligned}$$

Formulas (7) and (8) are more convenient for applications than those of § 3.

5. **Virtual Manifolds.**—The formulas developed so far do not always hold. Thus (2), which can be considered as the foundation of the rest, can be applied only when the dimensionality $r-h$ of M is >1, if its intersection with C_1C_2 has $r-h-2$ dimensions. It is the purpose of what follows to remove these restrictions. This will be done by the introduction, with Severi, of *virtual manifolds*.

If A, B, C are effective, and $A = B + C$, C is called the difference of A and B, and we write $C = A - B$. It is well known that $C = (A + D) - (B + D)$ for any effective D. Given that A, B are effective; if there

is no effective C such that $A = B + C$, we say that C is *virtual*. Similarly, if there is no effective $r - h$ dimensional manifold $M_{r-h} = C_1C_2 \cdots C_h$, we will say that M_{r-h} is virtual. This may happen only if the C's, though all effective, have in common a manifold of less than $r - h$ dimensions, or if some of the C's themselves are virtual. When M_{r-h} is virtual, it is a mere symbol; by definition it is unchanged when the C's are arbitrarily permuted.

6. Let A, B, C be effective and such that $C = A - B$, and also that BC^k be effective for $k \leqq r - 1$. By repeatedly applying (1) and (2), we obtain

$$\begin{aligned} \{A\} &= \{B + C + BC\}, \\ \{BA\} &= \{B^2 + BC + B^2C\}, \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \{B^{r-1}A\} &= \{B^r + B^{r-1}C + B^rC\}. \end{aligned}$$

Hence, by eliminating $BC, B^2C, \cdots, B^{r-1}C$,

$$\{C\} = \{(A - B)(1 - B + B^2 - \cdots + (-1)^{r-1}B^{r-1}\} + (-1)^r\{B^rC\};$$

$$\therefore \qquad \{C\} = \left[\frac{A - B}{1 + B}\right] + (-1)^r, \tag{9}$$

since $\{B^rC\} = +1$, (§ 2). In this formula and those to be developed, the fractions stand for their expansion in powers of $A, B, \cdots$. Formula (9) may be considered as the "inversion" of (1). Similarly if the manifolds

$$M = C_1C_2 \cdots C_h, \qquad MCB^k, \qquad k \leqq r - h - 1,$$

are all effective, the equation

$$\{MC\} = \left[M\frac{A - B}{1 + B}\right] + (-1)^{r-h} \tag{10}$$

gives the inversion of (2). If $C_i = A_i - B_i$, $(i = 1, 2, \cdots, h)$, and if the manifolds

$$\begin{aligned} &A_{i_1}A_{i_2} \cdots A_{i_k}B_1^{b_1}B_2^{b_2} \cdots B_h^{b_h}, \qquad (k + \Sigma b_i < r), \\ &C_1C_2 \cdots C_sB_1^{b_1}B_2^{b_2} \cdots B_h^{b_h}, \qquad (s \leqq h,\ s + \Sigma b_i < r), \end{aligned}$$

are all effective, the inversion of (6), obtained by repeated application of (10), is given by

$$\{C_1C_2 \cdots C_h\} = \left[\prod_{i=1}^{h}\frac{(A_i - B_i)}{1 + B_i}\right] + (-1)^{r+1-h}. \tag{11}$$

Suppose now that $M = C_1C_2 \cdots C_h$ be virtual. We can always find effective hypersurfaces A_i, B_i, such that $C_i = A_i - B_i$, then by definition

(a) $\{M\}$ is given by (11) as if M were effective.

(b) The relations between $\{M\}$ and $[M]$ are the same as when the manifold is effective.

If all manifolds $A_{i_1}A_{i_2}\cdots A_{i_k}B_1^{b_1}B_2^{b_2}\cdots B_h^{b_h}$, $k + \Sigma b_i < r$, are effective, this definition is complete. If not $[M]$ will be determined in terms of genera of manifolds, not all effective perhaps, but all situated in effective hypersurfaces, that is in a V_{r-1}, and the definition is then recurrent. In all cases a repeated application of (11) will lead by a finite number of operations to effective genera.

7. The genus of a manifold in the extended field composed of all effective and virtual manifolds has the following fundamental properties which are known to hold in the restricted field, and which justify the definitions just adopted.

I. *If* $C = A - B$, *where* A *and* B *are effective, then*

$$\{M \cdot C\} = \left[M \cdot \frac{A - B}{1 + B}\right] + (-1)^{r-h}.$$

This follows at once from (11) as applied to $C_1C_2\cdots C_hC$.

II. *If* $C = C' + C''$, *then*

$$\{MC\} = \{M((1 + C')(1 + C'') - 1)\}.$$

Let $C' = A' - B'$, $C'' = A'' - B''$, $C = \overline{A' + A''} - \overline{B' + B''}$, where A', B', A'', B'', are effective. The property to be proved holds for a V_1, for which, by definition, $[\overline{A - B}] = [A - B]$, so that the verification is immediate. Assume that it be true for any V_{r-1}, we say that it is also true for V_r. We have, by repeated application of property I,

$$\{M(C' + C'' + C'C'' - C)\} = \left[(C' + C'' + C'C'' - C)\prod_{i=1}^{h}\frac{A_i - B_i}{1 + B_i}\right].$$

If we expand the right-hand side, we obtain a sum of terms such as

$$[A_1 \cdot (C' + C'' + C'C'' - C)M'] - [B_1 \cdot (C' + C'' + C'C'' - C)M'].$$

As A_1 and B_1 are effective hypersurfaces, that is $r - 1$ dimensional varieties, it follows from our assumptions that the expression just written is equal to zero.

$$\therefore \quad \{M((1 + C')(1 + C'') - C - 1)\} = 0,$$

which is what we wished to prove.

Property II means that formula (2), and all formulas derived so far hold whether the manifolds involved are effective or not. As all the relations to be derived are based upon these, no restrictions in regard to "effectiveness" will be necessary.

III. *If in $C_i = A_i - B_i$, A_i and B_i be replaced by $A_i + D_i$, $B_i + D_i$, respectively,* $[M]$ *does not change.* In other words, $[M]$ is a function of the differences $A_i - B_i$. We have, by property II,

$$(-1)^{r+1-h} + \left[\prod_{i=1}^{h} \frac{\overline{A_i + D_i} - \overline{B_i + D_i}}{1 + \overline{B_i + D_i}}\right]$$

$$= \left[\prod_{i=1}^{h} \frac{(1 + A_i)(1 + D_i) - (1 + B_i)(1 + D_i)}{(1 + B_i)(1 + D_i)}\right] + (-1)^{r+1-h}$$

$$= \left[\prod_{i=1}^{h} \frac{A_i - B_i}{1 + B_i}\right] + (-1)^{r+1-h} = \{M\}$$

as was to be proved.

II. Solution of the fundamental problem.

8. In the solution of the fundamental problem, it is desirable to obtain formulas which hold, whatever the signs of the λ's as defined in § 1. The relations derived so far are not suitable for this purpose. We propose to show first the existence of solutions of the required nature. This having been done, the solutions will be obtained for a special case, and the generalization will follow easily.

Let $A = C_1 + C_2$. Whatever the positive integer λ, we have

$$\lambda A = \lambda C_1 + \lambda C_2.$$

Applying (7) to both sides of this relation, we obtain

$$[(1 + A)^\lambda - 1] - \binom{\lambda - 1}{r - 1} = [(1 + C_1)^\lambda(1 + C_2)^\lambda - 1] - \binom{2\lambda - 1}{r - 1}.$$

This equation of degree r in λ, has an infinity of roots, hence it is identically true. If we set

$$F(x) = \sum_{k=1}^{r} (-1)^{k-1}\frac{x^k}{k} + (-1)^{r-1}\sum_{k=1}^{r-1}\frac{1}{k},$$

we obtain by equating terms of the first degree in λ,

$$[F_r(A)] = [F_r(C_1)] + [F_r(C_2)].$$

This also shows that if $C_2 = A - C_1$, then

$$[F_r(C_2)] = [F_r(A)] - [F_r(C_1)].$$

Hence, if $\lambda C = \sum_{i=1}^{\rho}\lambda_i C_i$, then

$$\lambda[F_r(C)] = \left[\sum_{i=1}^{\rho}\lambda_i F_r(C_i)\right]. \tag{12}$$

Let $M = D_1 D_2 \cdots D_h$. If we reason from (2) as we just have from (1), and set

$$r - h = r'; \qquad F_{r'}(MC) = \sum_{k=1}^{r'} (-1)^{k-1} \frac{MC^k}{k} + (-1)^{r'-1} \sum_{k=1}^{r'-1} \frac{1}{k}$$

we can show that

$$\lambda[F_{r'}(MC)] = \left[\sum_{i=1}^{\rho} \lambda_i F_{r'}(MC_i)\right]. \tag{13}$$

Formulas (12) *and* (13) *hold whatever the signs of the* λ'*s.*

In the case $r = 2$, $[F_2(C)] = \frac{1}{2}[2C - C^2 - 2]$. The double of this last quantity has already been considered by Severi,[14] who proved (12) for $r = 2$, by considering the intersection of C with the canonic curves of V_2, and made this formula the central point in his solution of the fundamental problem for $r = 2$. But even before this, Castelnuovo and Enriques had shown that if $[F_2(C)] < 0$, for any continuous system of curves on the surface, the latter was birationally transformable into a scroll.

9. Let $(C_1, C_2, \cdots, C_\rho)$ form a base on V_r. For any C not belonging to the base, we have $\lambda C = \sum_{i=1}^{\rho} \lambda_i C_i$. Then

THEOREM. *The genus* $[C_i^h C^k]$ *is a linear function of the genera* $[C_1^{i_1} C_2^{i_2} \cdots C_\rho^{i_\rho}]$, *with coefficients equal to polynomials in the ratios* λ_i/λ, *the coefficients of these polynomials being independent of the signs of the* λ'*s.*

The proof will be by induction. The proposition is true: (a) If $h + k > r$, for then $[C_i^h C^k] = 0$. (b) If $h + k = s$, $k = 0$, that is for $[C_i^h]$. Now apply (13), making $M = C_i^{r-1}$. It gives

$$\lambda[CC_i^{r-1}] = \sum_{j=1}^{\rho} \lambda_j [C_j C_i^{r-1}],$$

which holds whatever the signs of the λ's, and shows that the theorem is true for $[CC_i^{r-1}]$, that is when $h + k = r$, $k = 1$. To prove it in the general case, assume that it holds if $h + k > s$, or if $h + k = s$, $k < t$. We say that it also holds when $h + k = s$, $k = t$. To show this, apply (10) with $M = C_i^{s-t} C^{t-1}$. It gives

$$\lambda \sum_{k=1}^{r-s+1} (-1)^{k-1} [C_i^{s-t} C^{t-1+k}] + (-1)^{r-s} \lambda \sum_{k=1}^{r-s} \frac{1}{k} = \sum_{j=1}^{\rho} \lambda_j [F_{r-s+1}(C_i^{s-t} C^{t-1} \cdot C_j)],$$

which also holds whatever the signs of the λ's. This relation, if solved for $[C_i^{s-t} C^t]$, shows that this genus is a linear combination of others of the type $[C_i^h C^k]$, such that $h + k > s$, or $h + k = s$, $k < t$, with coef-

[14] Sulla base per la totalita delle curve, l. c., p. 224.

ficients of the required character. Since for these last genera the theorem is true by assumption, it is also true for $[C_i^{s-t}C^t]$, which was to be proved.

In the same manner we may prove this more general theorem:

THEOREM. *If D_j is such that $\lambda_j D_j = \sum_i \lambda_i^j C_i$, then $[D_1 D_2 \cdots D_h]$ is a linear combination of the genera $[C_1^{i_1} C_2^{i_2} \cdots C_p^{i_p}]$, with coefficients polynomials in the ratios λ_i^j/λ_j, the coefficients of these polynomials being independent of the signs of the λ's.*

The two theorems just considered establish the existence of solutions of the fundamental problem, satisfying the condition to hold whatever the signs of the λ's.

Remark.—The question may be raised, whether the two preceding theorems hold when "division" is not unique,[15] that is when it is possible that $\lambda C = \lambda D$, while $C \neq D$ (λ integer > 1). This question must be answered in the affirmative. This will be proved, if we can show that if $\lambda C = \lambda D$, then $[MC^k] = [MD^k]$, whatever M, for then if $\lambda C = \Sigma \lambda_i C_i$, $[MC^k]$ will be determined without ambiguity. Let $M = A_1 A_2 \cdots A_h$. If $s + t > r - h$, then

$$[MC^s D^t] = [MC^{s+t}] = [MD^{s+t}] = 0.$$

Suppose that

$$[MC^s D^t] = [MC^{s+t}] = [MD^{s+t}], \qquad s + t > s_0,$$

we say that it is still true for $s + t = s_0$. For assume $s + t = s_0$, and apply (13). It follows that

$$\left[\sum_{k=1}^{(r-h+s_0)} (-1)^{k-1} \frac{MC^{s-1+k} D^t}{k}\right] = \left[\sum_{k=1}^{r-h+s_0} \frac{MC^{s-1} D^{t+k}}{k}\right].$$

The only terms which are not equal by assumption are those for which $k = 1$. Hence

$$[MC^s D^t] = [MC^{s-1} D^{t+1}] = [MC^{s-2} D^{t+2}] = \cdots = [MD^{s_0}] = [MC^{s_0}],$$

as was to be proved.

Definition.—If $[MC^k] = [MD^k]$, whatever the integer k, and the manifold M, the hypersurfaces C, D, are said to be arithmetically equivalent, and we write $C \sim D$. This equivalence has the following properties:

(a) In its field division is unique, that is if $\lambda C \sim \lambda D$, then $C \sim D$.

(b) If $C = D$, then $C \sim D$.

(c) If $C \sim D$, then $\lambda C = \lambda B$, where in general the integer $\lambda > 1$. For let $|H|$ be the system of hyperplane sections. The property is true on a V_1; for two groups of m points on an algebraic curve belong to the

[15] Severi, Ann. Ec. Norm. Sup., 1908, p. 456.

same continuous system. Grant that it holds on a V_{r-1}, we say that it also holds on V_r. We have by assumption $[HMC^k] = [HMD^k]$.

$$\therefore \quad HC \sim HD, \qquad \lambda HC = \lambda HD,$$

whatever the hyperplane section H. Hence[15'] $\lambda C = \lambda D$. In this, we have tacitly assumed that C, D are effective. In the contrary case, we would replace them by $C + kH$, $D + kH$, the integer k being so chosen that the two new hypersurfaces are effective. From $\lambda(C + kH) = \lambda(D + kH)$, follows then $\lambda C = \lambda D$. The properties here given show the relations between the two kinds of equivalence.

10. We are now ready to solve the fundamental problem. The reasoning used in § 9 could be utilized, this being the method followed by Severi for $r = 2$. It seems however impracticable for $r > 2$, and another will be followed here.

Let us specialize the hypersurface C of § 9, by making $\lambda = 1$, and assuming $\lambda_i \geqq 0$; $i = 1, 2, \cdots, \rho$. Then $C = \Sigma\lambda_i C_i$, and we can obtain $\{C\}$ by applying (7). C is the sum of $\Sigma\lambda_i$ hypersurfaces,

$$\therefore \quad \{C\} = \left[\prod_{i=1}^{\rho}(1 + C_i)^{\lambda_i} - 1\right] - \binom{\Sigma\lambda_i - 1}{r - 1}.$$

This result holds for all positive values of the λ's. From § 9, follows at once that if $\lambda C = \Sigma\lambda_i C_i$, and if we set $\mu = (1/\lambda)\Sigma\lambda_i$, then

$$\{C\} = \left[\prod_{i=1}^{\rho}(1 + C_i)^{\lambda_i/\lambda} - 1\right] - \binom{\mu - 1}{r - 1}. \tag{14}$$

Similarly we may prove the following result:

Consider the hypersurfaces defined by

$$\lambda_j D_j = \sum_i \lambda_i^j C_i \qquad (j = 1, 2, \cdots, h),$$

and set

$$\mu_j = (1/\lambda_j)\sum_i \lambda_i^j, \qquad \mu = \Sigma\mu_j,$$

then

$$\begin{aligned}\{D_1 D_2 \cdots D_h\} = \left[\prod_{j=1}^{h}\left(\prod_{i=1}^{\rho}(1 + C_i)^{\lambda_i^j/\lambda_j} - 1\right)\right] - \binom{\mu - 1}{r - 1} \\ + \sum_i \binom{\mu - \mu_i - 1}{r - 1} \cdots + (-1)^h \sum_i \binom{\mu_i - 1}{r - 1}\end{aligned} \tag{15}$$

and in particular

$$[D_1 D_2 \cdots D_r] = \left[\prod_{j=1}^{r}\left(\frac{1}{\lambda_j}\sum_{i=1}^{\rho}\lambda_i^j C_i\right)\right]. \tag{16}$$

In these formulas, the parentheses raised to fractional exponents, must be

[15'] Severi, l. c., p. 467.

expanded in series of ascending powers of the C's, and the result treated in accordance with the definition of the symbol [ ]. Formula (15) gives the most general solution of the fundamental problem.

Remark.—It is important to remember that these formulas hold whether the manifolds involved are effective or not, and whatever the signs of the λ's, that is *without any restrictions*. If $D_1D_2 \cdots D_h$ is virtual, its genus is the quantity defined in § 6. If it is effective, its arithmetic genus is the quantity defined in § 2. In particular (16) gives the number of points common to r hypersurfaces, in terms of their base characters, and is the immediate generalization of a formula of Severi's for $r = 2$.

III. Applications.

11. **Loci of Flats.**—A locus of flats is a V_r locus of $(r - 1)$-flats forming an irrational pencil. Let p be the genus of the pencil, H an arbitrary hyperplane section of V_r, G one of its generators (generating flats). It may be shown that H, G, form a *minimum base* for the variety.[16] That the latter can be considered as the projection of a non-singular variety from a higher space, will be admitted here, and the results obtained applied to this case. If $k > 1$, the intersection of H^i with G^k has less than $r - i - k$ dimensions.[17] As we have seen in § 5, H^iG^k must then be considered as virtual, and its genus cannot be determined from geometric considerations, but only by considering G as the difference of two appropriate effective hypersurfaces and applying (11). If we set $H_1 = H - G$, $H_1{}^iH^k$ is always effective, and clearly H, H_1 form also a minimum base—it will be the one used here. Before proceeding with direct applications, some preliminary propositions necessary to obtain the genera $[H_1{}^iH^k]$, will be proved.

12. Theorem. *If V_r, V_r' are two loci of flats, having in common a curve which meets the generators of each in one point only, they can be birationally transformed into each other.*

Let S_{r+k} be the space containing them both, P an arbitrary point of V_r, G the generator through P, m the point where this generator cuts the curve common to the two varieties. Through m there goes a generator G' of V_r'. If S_k is an arbitrary k-flat in S_{r+k}, the $(k + 1)$-flat determined by S_k and the point P, cuts G' in a single point P'. Conversely, given

---

[16] The proof can be given in various ways. For example by generalizing Severi's for $r = 2$. See: Sulle corrispondenze fra i punti di due curve e sopra certi classi di superficie, Memorie della R. Acc. delle Sc. di Torino, 1903, p. 22. Certain classical formulas of his and Segre's regarding scrolls, will also be found there.

[17] Unless V_r is a *hypercone*, that is such that its generators all go through the same S_{r-2}, called the *axis* of the hypercone. We will assume that this is not the case, though all results obtained hold also for a hypercone, when only variable intersections are considered.

P' on V_r', the point P on V_r, is determined in one and only one way. Clearly P, P' are in 1–1 correspondence, and this correspondence defines a birational transformation, which proves the theorem.

Corollaries. I. *If the two varieties have in common a d-dimensional manifold M, having an S_{d-1} for complete intersection with an arbitrary generator of either, V_r and V_r' can also be birationally transformed into each other.*

For M is clearly a locus of flats—the flats S_{d-1}—and a plane section of it, is a curve satisfying all conditions of the preceding theorem.

II. *If the conditions of the theorem, or corollary I, are satisfied then* $[V_r] = [V_r']$.

For two varieties which can be birationally transformed into each other, have the same arithmetic genus.[18]

13. Theorem. *V_r being the same locus of flats as before* $[V_r] = (-1)^{r-1} \cdot p$. This proposition was proved by Cayley[19] for the case $r = 2$, and for $r = 3$ it follows readily from a formula of Severi.[20]

Any V_r locus of $(r-1)$-flats can be considered as the hyperplane section of a V_{r+1} locus of r-flats. Its hyperplane section H is itself the locus of the $(r-2)$-flats HG, which form a pencil of genus p. It is therefore sufficient to prove that $H = (-1)^{r-2} \cdot p$. This will be done by induction. For $r = 2$, H is the plane section of a ruled surface and the proposition is trivial. We will assume that $r > 2$, that the theorem to be proved is true for a locus of flats of less than $r - 2$ dimensions, and prove its correctness for H.

We may suppose that V_r is immersed in an S_{r+1}, and this without any loss of generality, since we can always project it into such a space and reason upon the projection. Let Φ be an r-dimensional quadratic variety through a fixed generator G_0 of V_r, and call A its residual intersection with V_r. We say that A *is in one-one relation with* H, arbitrary hyperplane section of V_r. For let π be an arbitrary but fixed plane of G_0. When $r = 3$, π coincides with G_0. If V_r is not a hypercone, the generator G through any point P of the variety intersects G_0 into an S_{r-3}, which meets π in a single point m. This point is in Φ, hence the line mP cuts the quadratic variety in one point only, other than m, say P', a point which is in A. To P corresponds one point P'. Conversely, given P' arbitrary in A, m is determined as the point where G cuts π, and P is determined as the point where the line mP' cuts the hyperplane which intersects V_r into H. If V_r is a hypercone, we can take as point m, any

[18] Severi, Fundamenti. . . , p. 83.

[19] "On the deficiency of certain surfaces," Math. Ann., Vol. III, 1871, p. 526.

[20] Fundamenti. . . , p. 81.

point where Φ cuts the axis, and proceed as in the general case. In all cases P, P', and therefore also A and H, are in one-one correspondence.

It follows from this that $[A] = [H]$.

On the other hand, HH_1 cuts a generator HG of H, in a manifold HH_1G, which is an $(r - 3)$-flat, since H_1 is the residual intersection of V_r with an S_r through a generator. Furthermore H_1 is, like H, a locus of $(r - 2)$-flats, its generators being manifolds H_1G, each of which has in common with H_1 the same $r - 3$ flat HH_1G, which has just been considered.

$$\therefore \quad [H] = [H_1] \quad \text{(Coroll. II, § 12)}.$$

Also $A = H + H_1$, hence, by formula (1), since $r > 2$,

$$[A] = [H + H_1 + HH_1] = 2]H] + [HH_1] = [H];$$

$$\therefore \quad [H] = -[HH_1].$$

But HH_1 is the hyperplane section of H_1, and therefore a locus of $(r - 3)$-flats of a pencil of genus p, and by assumption the proposition is true for such a variety.

$$\therefore \quad [HH_1] = (-1)^{r-3} \cdot p, \qquad [H] = (-1)^{r-2}p,$$

which proves the theorem.

14. We are now in position to obtain the genera $[H^iH_1{}^k]$, which are necessary to solve the fundamental problem for V_r. The manifold $H^{i+1}H_1{}^k$ is an $r - i - k - 1$ space, if $i + k < r$. Hence

$$[H^iH_1{}^k] = (-1)^{r-i-k-1} \cdot p; \qquad (i + k < r).$$

Let μ be the order of V_r. Then $[H^r] = \mu$. To obtain $[H^{r-k}H_1{}^k]$, remark that applying formula (16), we have, since $[H^{r-k}H_1{}^{k-1}G] = 1$,

$$[H^{r-k}H_1{}^{k-1}(H_1 + G)] = [H^{r-k+1}H_1{}^{k-1}] = 1 + [H^{r-k}H_1{}^k];$$

$$\therefore \quad [H^{r-k}H_1{}^k] = [H^{r-k+1}H_1{}^{k-1} - 1].$$

Applying this last formula for all values of k, from 0 to k, and adding the results, we obtain

$$[H^{r-k}H_1{}^k] = \mu - k.$$

15. Let now C be any hypersurface in V_r. We have

$$C = \lambda H + \lambda_1 H_1$$

where λ, λ_1, are certain integers. Applying (14) we obtain

$$\{C\} = [(1 + H)^\lambda(1 + H_1)^{\lambda_1} - 1] - \binom{\lambda + \lambda_1 - 1}{r - 1}.$$

Expanding, and replacing the genera by their values as derived in § 14, we have, after some familiar transformations of combinatory analysis,

$$\{C\} = \binom{\lambda + \lambda_1 - 1}{r - 1}\left(p + \mu\frac{\lambda + \lambda_1}{r} - \lambda_1 - 1\right) + (-1)^r p.$$

Let a be the order of C, α that of its intersection CG with G. Then, by (16),

$$\alpha = [CH^{r-2}G] = [H^{r-2}(H - H_1)(\lambda H + \lambda_1 H_1)] = \lambda + \lambda_1,$$

$$a = [CH^{r-1}] = [(\lambda H + \lambda_1 H_1)H^{r-1}] = \mu\lambda + \lambda_1(\mu - 1) = \alpha\mu - \lambda_1.$$

Substituting in the formula obtained, we have finally

$$\{C\} = \binom{\alpha - 1}{r - 1}\left(p - 1 + \frac{\mu\alpha}{r} + a - \alpha\mu\right) + (-1)^r p. \qquad (17)$$

For $r = 2$, this formula becomes a well-known one of Segre,[21] namely

$$\{C\} = [C] = \alpha p - \mu\binom{\alpha}{2} + (a - 1)(\alpha - 1).$$

16. Let $M = C_1 C_2 \cdots C_h$, with $C_i = \lambda^i H + \lambda_1^i H_1$. To derive the genus of M, we might merely apply formula (15). However, it will be done more readily in the following way: A locus V_{r-h+1} of $(r - h)$-flats, which are in one-one correspondence with the generators of V_r, is of the same nature as V_r, and may be assumed to lie in the same S_{r+k}. Projecting the intersection of any generator of V_r with M, on the corresponding generator of V_{r-h+1}, a hypersurface M' of the latter variety is obtained. The orders of M and M', or of MG and $M'G'$, are the same—G' being a generator of V_{r-h+1}. Hence, it is sufficient to obtain the order a of M, and α of MG, and then to apply (17) on V_{r-h+1}, since $[M] = [M']$. Let a_i, α_i, be the order of C_i and $C_i G$, respectively. We have as previously,

$$\alpha_i = \lambda^i + \lambda_1^i, \qquad a_i = \mu\alpha_i - \lambda_1^i.$$

Hence, by (16),

$$a = \left[\prod_{i=1}^{h}(\lambda^i H + \lambda_1^i H_1)H^{r-h}\right] = \alpha\left(\mu + \sum_{i=1}^{h}\frac{a_i - \mu\alpha_i}{\alpha_i}\right),$$

$$\alpha = \left[\prod_{i=1}^{h}(\lambda^i H + \lambda_1^i H_1)H^{r-h-1}G\right] = \prod_{i=1}^{h}\alpha_i.$$

Substituting in (17), we have

$$\{M\} = \binom{\alpha - 1}{r - h}\left(a - \mu\alpha + p - 1 + \frac{\mu\alpha}{r - h + 1}\right) + (-1)^{r-h+1} \cdot p.$$

[21] "Courbes et surfaces réglées algébriques," Math. Ann., Vol. 34, 1890, p. 3.

For $h = r$ we have either from this last formula, or from (16),

$$[C_1C_2 \cdots C_r] = \left[\prod_{i=1}^{r}(\lambda^i H + \lambda_1{}^i H_1)\right] = \prod_{i=1}^{r}\alpha_i\left(\sum_{i=1}^{r}\frac{a_i}{\alpha_i} - (r-1)\mu\right),$$

which for $r = 2$, reduces to a well known relation on scrolls.

17. Complete intersections in S_r. In the case of an S_r, $\rho = 1$, and the minimum base is formed by an $(r-1)$-flat, say C. If m is the order of a V_r in S, we have $V_r = mC$. Also $[C^k] = 0$, if $k < r$, and $[C^r] = 1$. Hence if M is the complete intersection of h hypersurfaces of order m_1, $m_2, \cdots, m_h$, we obtain, by applying (15), when $h < r$ and with $m = \Sigma m_i$,

$$\begin{aligned}\{M\} = [M] &= \left[\prod_{i=1}^{h}(1+C)^{m_i} - 1\right] - \binom{m-1}{r-1} + \sum_{i=1}^{i=h}\binom{m-m_i-1}{r-1} \\ &\qquad + \cdots + (-1)^h\sum_{i=1}^{i=h}\binom{m_i-1}{r-1} \\ &= \binom{m-1}{r} - \sum_{i=1}^{i=h}\binom{m-m_i-1}{r} \\ &\qquad + \cdots + (-1)^{h-1}\sum_{i=1}^{i=h}\binom{m_i-1}{r}.\end{aligned}$$

The value of $[M]$ may be also obtained as follows: If $v(l)$ is the postulation of $[M]$, then

$$v(l) = \binom{l+r}{r} - \sum_i\binom{l-m_i+r}{r} + \sum_{i,j,i\neq j}\binom{l-m_i-m_j+r}{r} - \cdots.$$

Remembering that $[M] = (-1)^{r-h}\cdot(v(0)-1)$, we can from $v(l)$ obtain $[M]$ and, in fact, the same expression as the above is found. Conversely Severi's general expression[23] for $v(l)$ together with the formula derived here for $[M]$, can be utilized to obtain Bertini's result.

[22] Bertini, Introduzione alla geometria proiettiva degli iperspazi, p. 263.

[23] Fundamenti. . . , p. 41.

LAWRENCE, KANSAS,
January 6, 1916.

A CHARACTERISTIC PROPERTY OF SELF-PROJECTIVE CURVES.

By L. L. Dines.

A plane curve will be said to have the property P_{Ol} with respect to a point O and a line l of the plane, if the pencil of lines obtained by projecting a variable point of the curve from O is projective to the range of points cut out by the tangent at this variable point on l.

Evidently the ∞^3 conics through O and tangent to l have the property thus defined. It is the purpose of this paper to show that the property P_{Ol} characterizes the curves which admit infinitesimal projective transformations* into themselves. These curves have been known by various writers as self-projective curves, anharmonic curves, and W-curves.

The self-projective curves have many interesting properties which have been investigated by Klein,† Lie,‡ and others.§ The characteristic property P_{Ol}, which seems not to have been previously noted, furnishes for the self-projective curves a definition which is in some sense analogous to the Steinerian definition of conics.

1. The Differential Equations of Curves with Property P_{Ol}. In order to determine the curves having the property P_{Ol} (defined in our introduction) with respect to a fixed point O and line l, it is convenient to assume first that O is not on l, and to consider later the excepted case.

Case I. If O is not on l, we choose O as the vertex $(0, 0, 1)$ and l as the side $x_3 = 0$ of the triangle, with reference to which (x_1, x_2, x_3) shall represent the homogeneous coördinates of a point. We are to determine the curves

$$f(x_1, x_2, x_3) = 0 \tag{1}$$

with the property P_{Ol}.

Let P be any point of the curve (1) and let T be the intersection of the tangent at P with l. Then the curve (1) will have the property P_{Ol} if, and only if, the range $[T]$ is projective to the pencil $O[P]$, or what is equivalent, if the pencils $O[T]$ and $O[P]$ are projective. Since the equation

* An infinitesimal projective transformation is equivalent to a one-parameter group of finite projective transformations, and invariance under such a group may be used to define self-projective curves.

† Klein and Lie: "Ueber diejenigen Curven, welche durch ein geschlossenes System von unendlich vielen vertauschbaren Transformationen in sich uebergehen" (Math. Ann., IV, 1871).

‡ Lie-Scheffers: "Vorlesungen ueber continuierliche Gruppen."

§ For a brief treatment of these curves with bibliography, see Loria: "Spezielle algebraische und transcendente ebene Kurven," page 552.

of the tangent to (1) at $P(x_1, x_2, x_3)$ is $X_1 f_{x_1} + X_2 f_{x_2} + X_3 f_{x_3} = 0$, the coördinates of its intersection with the line $x_3 = 0$ are $(f_{x_2}, -f_{x_1}, 0)$, from which it follows that the coördinates of the line OT are $(f_{x_1}, f_{x_2}, 0)$. The coördinates of the line OP are easily seen to be $(x_2, -x_1, 0)$. The condition that the pencils $O[T]$ and $O[P]$ be projective is that a homogeneous bilinear relation exist between the coördinates of their corresponding elements. The most general relation of this sort is of the form

$$(2) \qquad (\alpha x_1 + \beta x_2) f_{x_1} + (\gamma x_1 + \delta x_2) f_{x_2} = 0,$$

where α, β, γ, and δ are arbitrary constants. Equation (2) is then the differential equation of the curves with property P_{ol}. It is desirable to reduce this equation, as well as others to follow, to non-homogeneous coördinates by means of the relations

$$(3) \qquad \frac{x_1}{x_3} = x, \qquad \frac{x_2}{x_3} = y.$$

The equation in non-homogeneous coördinates corresponding to the partial differential equation (2) is then equivalent to the ordinary differential equation

$$(2') \qquad \frac{dx}{\alpha x + \beta y} = \frac{dy}{\gamma x + \delta y}.$$

Since (2′) involves three independent parameters and is of first order, *it represents* ∞^4 *curves.*

Case II. If O is on l, we let l be the side $x_3 = 0$ and O the vertex $(1, 0, 0)$ of the reference triangle, and denote by O' the vertex $(0, 0, 1)$. Then if P and T have the same meanings as in the preceding case, the condition that the range $[T]$ and pencil $O[P]$ be projective is equivalent to the condition that the pencils $O'[T]$ and $O[P]$ be projective, and this latter projectivity finds analytic expression in the bilinear relation

$$(4) \qquad (\alpha x_2 + \beta x_3) f_{x_1} + (\gamma x_2 + \delta x_3) f_{x_2} = 0,$$

which upon reduction to non-homogeneous coördinates is equivalent to the ordinary differential equation

$$(4') \qquad \frac{dx}{\alpha y + \beta} = \frac{dy}{\gamma y + \delta}.$$

This equation like (2′) *represents* ∞^4 *curves* which have the property P_{ol}.

2. Concerning Self-Projective Curves. In his theory of infinitesimal transformations, Lie* considers the variation of a function due to an

* Lie-Scheffers: "Continuierliche Gruppen"; page 57. The author is indebted to Professor Eisenhart for suggestions regarding the use of Lie's results which led to a considerable shortening of the proof in this section.

infinitesimal projective transformation leaving invariant the line at infinity. Upon reduction to homogeneous coordinates by (3) with the assumption that x_3 remains constant, this variation takes the form

$$\delta f \equiv [(\alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3)f_{x_1} + (\alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3)f_{x_2}]\delta t,$$

where δt is an arbitrary infinitesimal, and α_{ij} are constants characterizing the transformation. The curves left invariant by this transformation are the integral curves of the differential equation

$$(\alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3)f_{x_1} + (\alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3)f_{x_2} = 0,$$

and are by definition *self-projective curves*. Since our equations (2) and (4) are of this form we have that *every curve with the property P_{Ol} is a self-projective curve.*

Lie has shown* that by suitable choice of a coördinate system, the differential equation representing the self-projective curves can always be reduced to one of the five canonical forms:

$$(5) \qquad (a)\ \frac{dx}{x} = \frac{dy}{ny}; \qquad (b)\ \frac{dx}{1} = \frac{dy}{y}; \qquad (c)\ \frac{dx}{y} = \frac{dy}{1}; \qquad (d)\ \frac{dx}{x} = \frac{dy}{y};$$
$$(e)\ \frac{dx}{0} = \frac{dy}{1}.$$

Evidently each of these equations is a special case of either (2′) or (4′). Hence *every self-projective curve has the property P_{Ol}*, and we have

THEOREM: *The property P_{Ol} is characteristic of self-projective curves.*

3. Types of Curves. Lie determines† the self-projective curves by integrating the canonical equations (5).

In the present connection it is of interest to examine the various types of curves with reference to the property P_{Ol}. For this purpose it is important to recall that the curves (2′) have the property P_{Ol} with respect to $O(0, 0, 1)$ and $l(x_3 = 0)$; while the curves (4′) have the property with respect to $O(1, 0, 0)$ and $l(x_3 = 0)$.

Type A. The integration of the canonical equation (a) gives a two parameter family of curves, each of which can however by a transformation, which changes only the unit point, be reduced to the form

$$(A) \qquad y = x^n.$$

Since (a) is a special case of (2′), each curve of the family (A) has the property P_{Ol} with respect to the vertex (0, 0, 1) and the opposite side $x_3 = 0$ of the reference triangle.

* Lie-Scheffers, loc. cit., page 64, Theorem 6; and page 69.

† Lie-Scheffers, loc. cit., pages 69–82.

A somewhat more general result is obtained by writing (A) in homogeneous coördinates in the form

$$x_1^{n_1}x_2^{n_2}x_3^{n_3} = 1,$$

where $n_1 + n_2 + n_3 = 0$. Since any transformation which merely interchanges the sides of the reference triangle transforms this equation into an equation of the same type, we may state more generally that every curve of the family (A) has the property P_{Ol} when O and l are respectively a vertex and opposite side of the reference triangle.

Type A'. The reduction of $(2')$ to (a) is accomplished by choosing as the reference triangle, the triangle which is invariant under that infinitesimal transformation for which curves (A) are self-projective. The vertex O and side l of this triangle may be supposed to be real. It may happen that the two sides of the invariant triangle which pass through O are imaginary. It is then advantageous to take these imaginary lines as the lines $y = \pm x\sqrt{-1}$. The equations of the curves represented by (A) may then be expressed in polar coördinates in the form

$$(A') \qquad \rho = e^{n'\theta}.$$

If l is the line at infinity, the curves (A') are the ordinary logarithmic spirals. That the logarithmic spiral has the property P_{Ol} with respect to the pole and the line at infinity, can be verified directly from the well known property of this curve that it cuts all its radii at a constant angle, which property is equivalent to the equality of the pencils $O[P]$ and $O[T]$.

Type B. The integration of (b) gives a family of curves each of which can be reduced, by a transformation which changes only the unit point, to the form

$$(B) \qquad y = e^x.$$

Since equation (b) is a special case of $(4')$, the curve (B) has the property P_{Ol} with respect to the point $(1, 0, 0)$ and the line $x_3 = 0$. But the curve (B) also has the property P_{Ol} with respect to the point $(0, 1, 0)$ and the line $x_2 = 0$. For (b) may be transformed into an equation of form $(2')$ by the transformation $x = x'/y'$, $y = 1/y'$; which changes $(0, 1, 0)$ into $(0, 0, 1)$, and $x_2 = 0$ into $x_3 = 0$.

If, in particular, $x_3 = 0$ is the line at infinity, (B) is the ordinary exponential curve. The exponential curve then has the property P_{Ol} with respect to the point at infinity on the x-axis, and the line at infinity; and also with respect to the point at infinity on the y-axis, and the x-axis.

Type C. Equation (c) is a special case of $(4')$; hence its integral curves

have the property p_{ol} with respect to the point (1, 0, 0) and the line $x_3 = 0$. Each integral curve can be reduced, by a transformation which does not change the coördinates of O or l, to the form

$$(C) \qquad y^2 = 2x.$$

This is a conic tangent to l at O, and the property P_{ol} is one of the classic properties of conics.

Types D and E. The integration of (d) and (e) gives the straight lines which obviously have the property P_{ol} with respect to any point and line.

4. Special case of conics. It was noted in our introduction that all conics through O and tangent to l have the property P_{ol}. If O is not on l, these conics are represented by equation (A) with $n = 2$. If O is on l, they are represented by (C).

Conics may also possess the property P_{ol} though they neither pass through O nor touch l. These are represented by (A) with $n = -1$ [or (A') with $n' = 0$], the point O being (0, 0, 1) and l being $x_3 = 0$. These are the conics with respect to which l is the polar of O. That such conics have the property P_{ol} can be verified directly from the fact that the lines OP and OT are in this case conjugate lines with respect to the conic; hence the pencils $O[P]$ and $O[T]$ are projective, the projectivity being in fact involutory.

University of Saskatchewan.

JUNE, 1916

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

PUBLISHED UNDER THE AUSPICES OF PRINCETON UNIVERSITY

SECOND SERIES, VOL. 17, NO. 4

LANCASTER, PA., AND PRINCETON, N. J.

CONTENTS

PAGE

On the Wronskian Test for Linear Dependence. By M. BÔCHER . . . 167

Note on a Theorem on Envelopes. By W. R. LONGLEY 169

Non-essential Singularities of Functions of Several Variables. By D. JACKSON 172

A Congruence of Circles. By F. W. BEAL 180

A Case of Iteration in Several Variables. By A. A. BENNETT . . . 188

The Arithmetic Genus of an Algebraic Manifold Immersed in Another. By S. F. LEFSCHETZ 197

A Characteristic Property of Self-Projective Curves. By L. L. DINES . 213

ANNALS OF MATHEMATICS

Published in September, December, March and June, at Lancaster, Pa. and Princeton, N. J., U. S. A.

Communications should be addressed to The Editors of the *Annals of Mathematics*, Princeton, N. J., U. S. A. Subscription price, $2 a volume (four numbers) in advance. Single numbers, 75 cents. All drafts and money orders should be made payable to THE ANNALS OF MATHEMATICS.

PRESS OF THE NEW ERA PRINTING COMPANY, LANCASTER, PA.